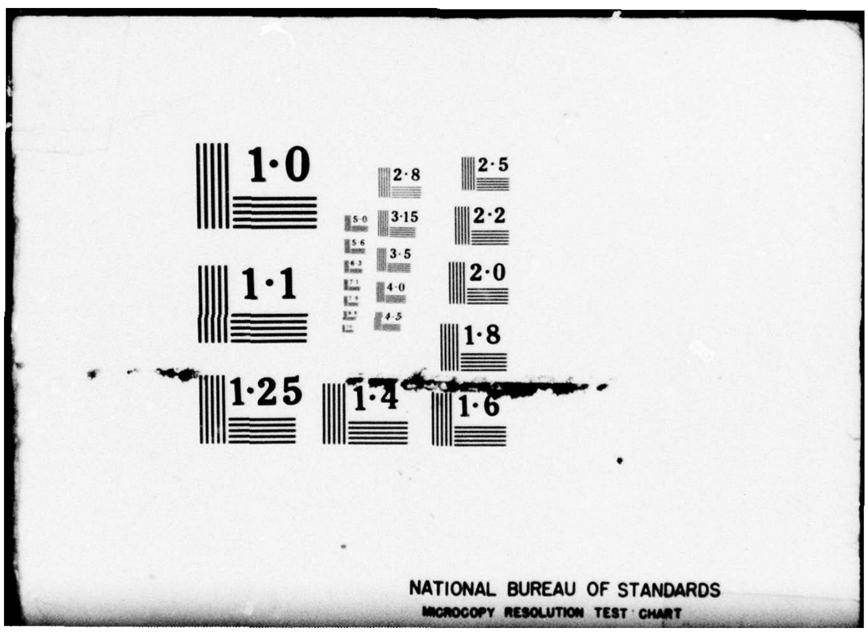
AD-A043 772 YALE UNIV NEW HAVEN CONN DEPT OF STATISTICS 12/1
ASYMPTOTIC DISTRIBUTIONS OF LATENT ROOTS IN CANONICAL--ETC(U)
JUN 77 W J GLYNN N00014-75-C-0563
UNCLASSIFIED TR-43 N/L

1 OF 2
ADA
043772


1·0
1·1
1·25
1·4
1·6
1·8
2·0
2·2
2·5
2·8
3·15
3·5
4·0
4·5
5·0
5·6
6·3
NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

12

YALE UNIVERSITY

Department of Statistics

Technical Report No. 43

TR-43

June, 1977

"Asymptotic distributions of latent roots in canonical correlation analysis and in discriminant analysis with applications to testing and estimation."

By William J. Glynn

147 p.

This research was supported by the Army, Navy, Air Force and NASA under a contract administered by the Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Contract number: N00014-75-C-0563

Task number: NR 042-242

402200

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1. An Overview.

The commonly used multivariate techniques of principal components, factor analysis, canonical analysis, multivariate analysis of variance, and discriminant analysis are based on latent roots and latent vectors of random matrices. Unfortunately, even under the usual assumption that the observations have been drawn from a multivariate normal population, the exact sampling distributions of the sample latent roots and sample latent vectors are generally so complicated that they are practically useless for computational or inferential purposes. The probability density functions, and hence the likelihood functions, are very difficult to evaluate numerically and are analytically intractable. Also it is not at all clear how one should draw inferences about a subset of the population parameters, the remaining population parameters being nuisance parameters. Asymptotic approximations to the exact sampling distributions appear to be necessary and have produced some useful results (see for example the work of Anderson (1965) and James (1969) in principal components).

This research considers sample latent roots, calculated from a sample from a multivariate normal population, which arise in

(i) canonical correlation analysis, and in

(ii) multivariate analysis of variance (MANOVA) and discriminant analysis.

The exact probability density functions (pdf's) of the sample roots in both of these cases involve hypergeometric functions of matrix arguments. Constantine (1963) found a representation of these hypergeometric func-

tions as a series of zonal polynomials, but in general the zonal polynomials are very difficult to calculate and the series converge very slowly. Asymptotic expansions of the hypergeometric functions are derived and these expansions are then used to obtain expansions of the pdf's, and hence of the likelihood functions. The resulting expansions are used to study inference problems. In canonical correlation analysis, estimates of the population coefficients are obtained and the Bartlett-Lawley tests that the residual population roots are zero are examined. In MANOVA and discriminant analysis, Bartlett's test that the residual population roots are zero is examined.

The rest of this chapter contains a review of some basic facts about zonal polynomials and hypergeometric functions of matrix arguments, a review of the main distribution and testing results in canonical correlations and MANOVA and discriminant analysis, and a brief summary of the remaining chapters and notation.

1.2. Zonal Polynomials and Hypergeometric Functions of Matrix Arguments.

Let S and T be $m \times m$ complex symmetric matrices and let $a_1,\ldots,a_p,b_1,\ldots,b_q$ be complex constants. The hypergeometric functions ${}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;S)$ and ${}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;S,T)$ are complex valued symmetric functions of the latent roots of S and S and T respectively. When the values of the a_i's and b_j's are clear from the context, we will denote ${}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;S)$ by ${}_pF_q(S)$ and ${}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;S,T)$ by ${}_pF_q(S,T)$. Also, when it is important to emphasize the dimension of S and T we will write ${}_pF_q^{(m)}(a_1,\ldots,a_p;b_1,\ldots,b_q;S)$ (or ${}_pF_q^{(m)}(S)$) and ${}_pF_q^{(m)}(a_1,\ldots,a_p;b_1,\ldots,b_q;S,T)$ (or ${}_pF_q^{(m)}(S,T)$).

Herz (1955) developed the general system of hypergeometric functions of one matrix argument by means of multidimensional Laplace and inverse Laplace transforms. Constantine (1963) found a representation of hypergeometric functions as a series involving zonal polynomials. We take Constantine's series representation as our definition.

Definition --

$$ {}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;S) = \sum_{k=0}^{\infty} \sum_{\varkappa} \frac{(a_1)_\varkappa \cdots (a_p)_\varkappa}{(b_1)_\varkappa \cdots (b_q)_\varkappa} \frac{C_\varkappa(S)}{k!} \tag{1.2.1} $$

where

$$ (a)_\varkappa = \prod_{i=1}^{m} (a-\tfrac{1}{2}(i-1))_{k_i} $$

and

$$ (a)_k = a(a+1)\cdots(a+k-1) \; . $$

$C_\varkappa(S)$ denotes the zonal polynomial of S corresponding to the partition $\varkappa = (k_1,\ldots,k_m)$, $k_1 \geq k_2 \geq \cdots \geq k_m \geq 0$ and $k_1 + k_2 + \cdots + k_m = k$, of k into not more than m parts. $C_\varkappa(S)$ is a symmetric homogeneous polynomial of degree k in the latent roots of S . A detailed discussion of zonal polynomials is presented in James (1961,1964) and Constantine (196[illegible]

If $p > q + 1$, the series [illegible]2.1) may only converge for $S = 0$. If $p = q + 1$, the series converges absolutely for $|S| < 1$ where $|S|$ denotes the maximum of the absolute values of the latent roots of S . If $p \leq q$, the series converges for all S . The constants a_i , $i=1,\ldots,p$, and b_j , $j=1,\ldots,q$, are arbitrary except that the b_j's must not be integers or half integers $\leq \frac{1}{2}(m-1)$.

James (1964) extended (1.2.1) to define hypergeometric functions of two matrix arguments by:

Definition --

$$(1.2.2)\quad {}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;S,T) = \sum_{k=0}^{\infty}\sum_{\kappa}\frac{(a_1)_\kappa\cdots(a_p)_\kappa}{(b_1)_\kappa\cdots(b_q)_\kappa}\frac{C_\kappa(S)C_\kappa(T)}{C_\kappa(I_m)k!}$$

where I_m is the $m \times m$ identity matrix.

Zonal polynomials and hypergeometric functions of products ST of symmetric matrices are defined as the corresponding zonal polynomials or hypergeometric functions of the symmetric matrix $S^{\frac{1}{2}}TS^{\frac{1}{2}}$ or $T^{\frac{1}{2}}ST^{\frac{1}{2}}$. Note that ST, TS, $T^{\frac{1}{2}}ST^{\frac{1}{2}}$, and $S^{\frac{1}{2}}TS^{\frac{1}{2}}$ have the same latent roots.

James (1961) proved that

$${}_0F_0(S) = \sum_{k=0}^{\infty}\sum_{\kappa}\frac{C_\kappa(S)}{k!} = \mathrm{etr}(S)$$

where

$$\mathrm{etr}(S) \equiv \exp(\mathrm{trace}\ S)\ .$$

Herz (1955) derived the general system of hypergeometric functions from ${}_0F_0(S)$ by means of multidimensional Laplace and inverse Laplace transforms. Let $f(S)$ be a function of the $m \times m$ positive definite symmetric matrix S. The Laplace transform $g(Z)$ of $f(S)$ is defined as

$$g(Z) = \int_{S>0} \mathrm{etr}(-SZ)f(S)(dS)$$

where $Z = X + iY$ is an $m \times m$ complex symmetric matrix, X and Y are real symmetric matrices, and the integral is assumed to converge in the half plane $\mathrm{Re}(Z) = X > X_0$ for some positive definite X_0. The notation $X > X_0$ means that $X - X_0$ is positive definite. The inverse Laplace transform is defined by

$$f(S) = \frac{2^{\frac{1}{2}m(m-1)}}{(2\pi i)^{\frac{1}{2}m(m+1)}}\int_{\mathrm{Re}(Z)>X_0>0} \mathrm{etr}(SZ)g(Z)(dZ)\ .$$

The integral is taken over $Z = X + iY$ with X_0 fixed, $X > X_0$, and Y varying over all real symmetric matrices. See Herz (1955) for details.

Herz proved the following:

$$(1.2.3)\quad {}_{p+1}F_q(a_1,\ldots,a_p,a;b_1,\ldots,b_q;S) = \frac{1}{\Gamma_m(a)}\int_{T>0} \operatorname{etr}(-T)\det T^{a-\frac{1}{2}(m+1)}\, {}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;ST)(dT)$$

when $a > \frac{1}{2}(m-1)$, where

$$(1.2.4)\quad \Gamma_m(a) = \pi^{m(m-1)/4}\prod_{i=1}^{m}\Gamma(a-\tfrac{1}{2}(i-1))$$

is the multivariate gamma function; and

$${}_pF_{q+1}(a_1,\ldots,a_p;b_1,\ldots,b_q,b;S) = \frac{2^{\frac{1}{2}m(m-1)}\Gamma_m(b)}{(2\pi i)^{\frac{1}{2}m(m+1)}}\int_{\operatorname{Re}(T)=X_0>0} \operatorname{etr}(T)\det T^{-b}\, {}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;T^{-1}S)(dT) .$$

We will need the following results which were proved by James (1964).

1. If $S = \operatorname{diag}(S_1,0)$ where S is $m\times m$, S_1 is a $k\times k$ complex symmetric matrix, and 0 is the $(m-k)\times(m-k)$ zero matrix, then

$$(1.2.5)\quad {}_pF_q^{(m)}(a_1,\ldots,a_p;b_1,\ldots,b_q;S) = {}_pF_q^{(k)}(a_1,\ldots,a_p;b_1,\ldots,b_q;S_1) .$$

2. Let $O(m)$ denote the group of orthogonal $m\times m$ matrices and let (dH) be the invariant measure on $O(m)$ normalized so that

$$\int_{O(m)} (dH) = 1 .$$

Then

$$(1.2.6)\quad {}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;S,T) = \int_{O(m)} {}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;SH'TH)(dH) .$$

3. Bessel's Integral --

Let X be an $\ell \times m$ matrix, $\ell \leq m$, and let $H = [H_1 : H_2] \in O(m)$ where H_1 is $m \times \ell$ and H_2 is $m \times (m-\ell)$. Then

$$\int_{O(m)} \operatorname{etr}(XH_1)(dH) = {}_0F_1(\tfrac{1}{2}m; \tfrac{1}{4}XX') \tag{1.2.7}$$

A detailed discussion of hypergeometric functions of matrix argument can be found in Herz (1955), Constantine (1963), and James (1964).

1.3. Canonical Correlation Coefficients.

Canonical correlations were introduced by Hotelling (1936) as a means of measuring the relation between two sets of variates. Let x be a $p \times 1$ random vector with mean μ and covariance matrix Σ_{11}. Let y be a $q \times 1$ random vector $(q \geq p)$ with mean λ and covariance matrix Σ_{22}. Let Σ_{12} $(p \times q)$ be the matrix of covariances between x and y. Hotelling proved that there exist nonsingular matrices A $(p \times p)$ and B $(q \times q)$ such that if $u = Ax$ and $v = By$ then the covariance matrix of the random $(p+q)$-vector $\binom{u}{v}$ is

$$\begin{bmatrix} I_p & \vdots & P & \vdots & 0 \\ P & \vdots & I_p & \vdots & 0 \\ 0 & \vdots & 0 & \vdots & I_{q-p} \end{bmatrix}$$

where $P = \operatorname{diag}(\rho_1, \ldots, \rho_p)$ is a diagonal matrix with $1 \geq \rho_1 \geq \rho_2 \geq \cdots \geq \rho_p \geq 0$. The ρ_i's are the canonical correlation coefficients. It is easily shown that the canonical correlation coefficients satisfy the determinantal equation

$$\det(\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - \rho^2\Sigma_{11}) = 0 .$$

Let $X = [x^1,\ldots,x^N]$ and $Y = [y^1,\ldots,y^N]$, where $x^1,\ldots,x^N$ and $y^1,\ldots,y^N$ are random samples of size $N = n+1$ from x and y respectively. The matrix S of corrected sums of squares and sums of products (abbreviated as corrected s.s. and s.p. matrix) of $[X' \vdots Y']'$, is defined as

$$S = \begin{bmatrix} X \\ Y \end{bmatrix} [I_N - N^{-1}E_{NN}] [X' \vdots Y']$$

where E_{ab} denotes an $a \times b$ matrix of ones. Partition S as follows

$$S = \begin{array}{cc} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} & \begin{array}{c} p \\ q \end{array} \\ \begin{array}{cc} p \quad & \quad q \end{array} & \end{array}, \qquad S_{12} = S_{21}'$$

where S_{11} is the corrected s.s. and s.p. matrix of X, S_{22} is the corrected s.s. and s.p. matrix of Y, and S_{12} is the corrected s.p. matrix of X and Y.

Assume that x and y have a joint multivariate normal distribution. Then the maximum likelihood estimates $1 \geq r_1^2 \geq r_2^2 \geq \cdots \geq r_p^2 \geq 0$ of the ρ_i^2's are the roots of the determinantal equation

$$\det(S_{12}S_{22}^{-1}S_{21} - r^2 S_{11}) = 0 .$$

Note that if the ρ_i's are strictly less than 1 then the r_i^2 are distinct and between 0 and 1 with probability 1.

In the following discussion x and y are assumed to have a joint normal distribution.

Fisher (1939), Roy (1939), and Hsu (1939) found the joint null distribution, corresponding to $P = 0$, of $r_1^2,\ldots,r_p^2$. Hsu (1941b) found the limiting joint pdf as $n \to \infty$ of $z_1,\ldots,z_p$ where

$$z_i = \frac{n^{\frac{1}{2}}(r_i^2 - \rho_i^2)}{2\rho_i(1-\rho_i^2)} \qquad \text{if} \quad \rho_i > 0$$

and

$$z_i = nr_i^2 \quad \text{if} \quad \rho_i = 0 .$$

Asymptotically the z_i's corresponding to different ρ_i's are independent. The z_i's corresponding to equal nonzero ρ_i's have a joint "normal type" limiting distribution while the z_i's corresponding to $\rho_i = 0$ have a joint "chi-square type" limiting distribution. Constantine (1963) derived the exact joint density of $r_1^2, \ldots, r_p^2$ when $n \geq p + q$. Let $P = \mathrm{diag}(\rho_1, \ldots, \rho_p)$ and $R = \mathrm{diag}(r_1, \ldots, r_p)$, then $f(R^2)$, the joint pdf of $r_1^2, \ldots, r_p^2$, is

$$(1.3.1)\quad f(R^2) = \prod_{i=1}^{p} (1-\rho_i^2)^{\frac{1}{2}n} \, {}_2F_1(\tfrac{1}{2}n, \tfrac{1}{2}n; \tfrac{1}{2}q; P^2, R^2)$$

$$\times C_1 \prod_{i=1}^{p} \{(r_i^2)^{\frac{1}{2}(q-p-1)} (1-r_i^2)^{\frac{1}{2}(n-q-p-1)}\} \prod_{i<j}^{p} (r_i^2 - r_j^2) \prod_{i=1}^{p} dr_i^2$$

$$1 \geq r_1^2 \geq r_2^2 \geq \cdots \geq r_p^2 \geq 0$$

where

$$C_1 = \frac{\Gamma_p(\frac{1}{2}n)\pi^{\frac{1}{2}p^2}}{\Gamma_p\{\frac{1}{2}(n-q)\}\Gamma_p(\frac{1}{2}q)\Gamma_p(\frac{1}{2}p)}$$

and Γ_p is the multivariate gamma function defined by (1.2.4). Chattopadhyay and Pillai (1973) and Chattopadhyay, Pillai, and Li (1976) found an asymptotic expansion as $n \to \infty$ of the ${}_2F_1$ hypergeometric function in (1.3.1). Unfortunately their result involves a ${}_2F_1$ function with the matrix P^2R^2 as argument. Sugiura (1976), using a perturbation argument, calculated terms up to and including terms of order n^{-1} in the joint pdf $z_1, \ldots, z_p$. If the terms of order n^{-1} are ignored, then the z_i's corresponding to distinct ρ_i's are independent.

Typically in canonical correlation analysis one is initially interested in testing the null hypothesis H_0 that all the population canonical correlation coefficients are zero versus the alternative that they are not all zero. H_0 is equivalent to the null hypothesis that x and y are independent. The likelihood ratio statistic L for testing H_0 is defined by

$$L^{\frac{1}{2}N} = \prod_{i=1}^{p} (1-r_i^2) = T_0 .$$

Bartlett (1938) showed that $-\{n-\frac{1}{2}(p+q+1)\}\ln T_0$ (where $n = N - 1$) is approximately distributed as χ^2 with pq degrees of freedom (df). The factor $n - \frac{1}{2}(p+q+1)$ was chosen so that the resulting statistic has the same moments up to order N^{-2} as χ^2 on pq df. An asymptotic expansion for the distribution function of Bartlett's statistic has been given by Box (1949) and Anderson (1958, Chapter 9).

If H_0 is rejected then x and y are correlated and often one would be interested in summarizing the relationship between x and y by means of a few canonical variables. Bartlett (1938, 1941, 1947a) suggested that the null hypothesis H_k, that the residual $p-k$ canonical correlation coefficients are zero when the largest k population coefficients, corresponding to real relationships between x and y, have been removed, can be tested with the statistic $\{n - \frac{1}{2}(p+q+1)\}T_k$ where

$$T_k = -\ln \prod_{i=k+1}^{p} (1-r_i^2) .$$

When H_k is true $\{n - \frac{1}{2}(p+q+1)\}T_k$ is approximately distributed as χ^2 with $(p-k)(q-k)$ df. If H_k is accepted then the canonical variables corresponding to $\rho_{k+1},\ldots,\rho_p$ have no predictive value and the relationship between x and y may be summarized by means of the first k

canonical variables. In this test the k largest population correlation coefficients are nuisance parameters. Lawley (1959) examined their effect on T_k by a complicated expansion of the matrices involved in the calculation of the r_i's. He concluded that for large n

$$\{n - k - \tfrac{1}{2}(p+q+1) + \sum_{i=1}^{k} \rho_i^{-2}\}T_k$$

is approximately χ^2 with $(p-k)(q-k)$ df. The correction factor was chosen so that the resulting statistic has the same moments up to order N^{-2} as χ^2 with $(p-k)(q-k)$ df. Since in practice one does not know the values of the k largest population canonical correlation coefficients, Lawley suggested that the statistic

$$\{n - k - \tfrac{1}{2}(p+q+1) + \sum_{i=1}^{k} r_i^{-2}\}T_k$$

should be used. Little appears to be known about the accuracy of such an approximation.

1.4. Multivariate Analysis of Variance and Discriminant Analysis.

Let $x^1, \ldots, x^\ell$ be $\ell = n_1 + 1$ independent p-variate $(n_1 \geq p)$ normally distributed random vectors such that x^i has mean μ^i and covariance matrix Σ, $i=1,\ldots,\ell$. Suppose one wants to test the null hypothesis

$$H_0 : \mu^1 = \mu^2 = \cdots = \mu^\ell$$

of the equality of the ℓ mean vectors on the basis of ℓ independent random samples of sizes m_i $(i=1,2,\ldots,\ell)$ from the ℓ populations. Define $M = m_1 + m_2 + \cdots + m_\ell$. Let X_i be the $p \times m_i$ matrix of the m_i sample observations from x^i. The corrected s.s. and s.p. matrix S_i of X_i is

$$S_i = X_i(I_{m_i} - m_i^{-1}E_{m_i m_i})X_i' \qquad i=1,\ldots,\ell$$

where E_{ab} is an $a \times b$ matrix of ones. The vector of sample means of observations from x^i is

$$\bar{x}^i = m_i^{-1}X_iE_{m_i 1}, \qquad i=1,\ldots,\ell .$$

The "within groups" matrix of s.s. and s.p. W is defined as $W = S_1 + \cdots + S_\ell$. W has a Wishart distribution on $n_2 = M - \ell$ df and associated matrix Σ (denoted as $W_p(n_2|\Sigma)$). The "between groups" matrix of s.s. and s.p. B is defined as

$$B = \sum_{i=1}^{\ell} m_i(\bar{x}^i - \bar{x})(\bar{x}^i - \bar{x})'$$

where

$$\bar{x} = M^{-1}\sum_{i=1}^{\ell} m_i\bar{x}^i .$$

B has a noncentral Wishart distribution on n_1 df with associated matrix Σ and noncentrality matrix $\Sigma^{-1}\Lambda$ (denoted as $W_p(n_1|\Sigma|\Sigma^{-1}\Lambda)$), where

$$\Lambda = \sum_{i=1}^{\ell} m_i(\mu^i - \mu)(\mu^i - \mu)'$$

and

$$\mu = M^{-1}\sum_{i=1}^{\ell} m_i\mu^i .$$

H_0 is equivalent to the hypothesis that $\Lambda = 0$. The likelihood ratio statistic for testing H_0 (Wilks (1932)) is

$$\det\{B(B+W)^{-1}\}^{\frac{1}{2}M} = \{\prod_{i=1}^{p}(1-\ell_i)\}^{\frac{1}{2}M}$$

where $1 \geq \ell_1 \geq \ell_2 \geq \cdots \geq \ell_p \geq 0$ are the latent roots of $B(B+W)^{-1}$. Note that with probability 1, $1 > \ell_1 > \ell_2 > \cdots > \ell_p > 0$.

Fisher (1939), Hsu (1939), and Roy (1939) found the joint null distribution, defined by $\Lambda = 0$, of $\ell_1, \ldots, \ell_p$. Constantine (1963) derived the exact joint density. Let $L = \mathrm{diag}(\ell_1, \ldots, \ell_p)$ and $\Omega = \mathrm{diag}(\omega_1, \ldots, \omega_p)$ where $\omega_1 \geq \omega_2 \geq \cdots \geq \omega_p \geq 0$ are the latent roots of $\Sigma^{-1}\Lambda$. The joint pdf of $\ell_1, \ldots, \ell_p$, $f(L)$, is

$$f(L) = \mathrm{etr}(-\tfrac{1}{2}\Omega)\, {}_1F_1\{\tfrac{1}{2}(n_1+n_2); \tfrac{1}{2}n_1; \tfrac{1}{2}\Omega; L)$$

$$\times\, C_1 \prod_{i=1}^{p} \{\ell_i^{\frac{1}{2}(n_1-p-1)} (1-\ell_i)^{\frac{1}{2}(n_2-p-1)}\} \prod_{i<j}^{p} (\ell_i - \ell_j) \prod_{i=1}^{p} d\ell_i \tag{1.4.1}$$

$$1 \geq \ell_1 \geq \ell_2 \geq \cdots \geq \ell_p \geq 0$$

where

$$C_1 = \frac{\pi^{\frac{1}{2}p^2}\, \Gamma_p\{\tfrac{1}{2}(n_1+n_2)\}}{\Gamma_p(\tfrac{1}{2}n_1)\Gamma_p(\tfrac{1}{2}n_2)\Gamma_p(\tfrac{1}{2}p)} .$$

Asymptotic approximations to the distribution of L (or functions of L) have been developed for either large n_2 (large error df) or for large Ω (some or all of the ω_i's large). If $\Lambda \neq 0$ then in most situations as n_2, and hence M, become large, Λ, and hence Ω, will become large. In particular, if the sample sizes $m_1, \ldots, m_\ell$ are proportional to M, then Ω will be proportional to M. This means that a reasonable asymptotic approach is to let $\Omega = n_2\Theta$ where $\Theta = \mathrm{diag}(\theta_1, \ldots, \theta_p)$ and let $n_2 \to \infty$. Hsu (1941a) found the limiting joint pdf as $n_2 \to \infty$ with $\Omega = n_2\Theta$ of $z_1, \ldots, z_p$, defined by

$$z_i = \sigma_i^{-1} n_2^{\frac{1}{2}} \{\ell_i(1-\ell_i)^{-1} - \theta_i\} \qquad \text{if} \quad \theta_i > 0$$

where $\sigma_i = (2\theta_i)^{\frac{1}{2}}(\theta_i+2)^{\frac{1}{2}}$, and

$$z_i = n_2\ell_i(1-\ell_i)^{-1} \qquad \text{if} \quad \theta_i = 0 .$$

Asymptotically the z_i's corresponding to different θ_i's are independent. The z_i's corresponding to equal nonzero θ_i's have a joint "normal type" limiting distribution; the z_i's corresponding to $\theta_i = 0$ have a joint "chi-square type" distribution. Fujikoshi (1976) extended Hsu's results using a perturbation argument. Chattopadhyay and Pillai (1973) and Chattopadhyay, Pillai, and Li (1976) found an asymptotic expansion as $n_2 \to \infty$ of the ${}_1F_1$ hypergeometric function in (1.4.1). Unfortunately their result involves another ${}_1F_1$ function with the matrix $-\frac{1}{2}L\Omega$ as argument. Constantine and Muirhead (1976) found an asymptotic expansion of the ${}_1F_1$ function in (1.4.1) when some or all of the ω_i are large.

In a typical multivariate analysis of variance situation one would usually be interested in testing H_0, at least as a first step. H_0 is equivalent to the null hypothesis that $\Omega = 0$. The likelihood ratio statistic L for testing H_0 is defined by

$$L^{\frac{1}{2}M} = \prod_{i=1}^{p} (1-\ell_i) = T_0 .$$

Bartlett (1938) showed that $-[n_2+\frac{1}{2}(n_1-p-1)]\ln T_0$ is approximately distributed as χ^2 with pn_1 df. The factor $n_2 + \frac{1}{2}(n_1-p-1)$ was chosen so that the resulting statistic has the same moments up to order M^{-2} as χ^2 with pn_1 df. An asymptotic expansion for the distribution function of Bartlett's statistic has been given by Box (1949) and Anderson (1958, Chapter 9).

If H_0 is rejected then the population means are not all equal and hence it is reasonable to look for linear functions which best discriminate among the populations. The number of meaningful discriminant functions is equal to the dimension of the subspace spanned by the population means,

or equivalently, to the rank of the noncentrality matrix which is equal to the number of nonzero ω_i . Bartlett (1947a) suggested that the null hypothesis H_k , that the population means lie in a subspace of dimension $p - k$, can be tested with the statistic

$$L_k = -\{n_2 + \tfrac{1}{2}(n_1-p-1)\}\ln \prod_{i=k+1}^{p} (1-\ell_i) .$$

When H_k is true L_k is approximately distributed as χ^2 with $(p-k)(n_1-k)$ df. H_k is equivalent to the hypothesis that the residual $p - k$ latent roots of the noncentrality matrix are zero given that the largest k latent roots are nonzero.

1.5. Summary of Results.

In Chapter 2 "asymptotic expansion" and "asymptotic representation" are defined and the technique which will be used to derive expansions is developed.

In Chapter 3 the partial differential equations, derived by Constantine and Muirhead (1972) for the ${}_2F_1$, ${}_1F_1$, ${}_0F_1$, ${}_1F_0$, and ${}_0F_0$ hypergeometric functions of two $p \times p$ symmetric matrices S and T are generalized to include the case where the smallest $p - k$ latent roots of T are known to be zero. The technique presented in Chapter 2 is then used to obtain an asymptotic expansion up to and including terms of order n^{-1} , of the ${}_2F_1$ function which occurs in the pdf of the squared sample canonical correlation coefficients. The nonzero population coefficients are assumed to be distinct.

In Chapter 4 the results of Chapter 3 are used to obtain a "beta-type" asymptotic approximation for large n of the joint pdf of the sample canonical correlation coefficients. In contrast to the more highly asymp-

totic "normal-type" approximations of Hsu (1941b) and Sugiura (1976), the sample roots are not independent up to and including terms of order n^{-1}. The "beta-type" approximation serves as a basis for the study of the estimation of the population coefficients and for the study of the Bartlett-Lawley tests that the residual population canonical correlation coefficients are zero. Since the distribution of the sample coefficients depends only on the population coefficients, the part of the distribution involving the population coefficients may be regarded as a marginal likelihood function. It is shown that Fisher's z-transformation applied to the marginal maximum likelihood estimates of the population coefficients gives unbiased estimates to order n^{-2} of the corresponding transformed population coefficients. These estimates have variance equal to $n^{-1} + O(n^{-2})$. The Bartlett-Lawley tests are investigated using an approach which is similar to that used by James (1969) in connection with tests of equality of the latent roots of a covariance matrix. The largest k sample coefficients are asymptotically sufficient for the k largest population coefficients and hence the asymptotic conditional distribution of the $p - k$ smallest sample coefficients, given the k largest sample coefficients, does not depend on any unknown parameters. The conditional distribution is used to confirm Lawley's correction factor and provide some information on the accuracy of the χ^2 approximation.

In Chapter 5 an asymptotic expansion is obtained, up to and including terms of n_2^{-1}, for large n_2 and $\Sigma^{-1}\Delta = n_2\Theta$, of the ${}_1F_1$ hypergeometric function which occurs in MANOVA and discriminant analysis. The nonzero latent roots of the noncentrality matrix $\Sigma^{-1}\Delta$ are assumed to be simple.

In Chapter 6 the results of Chapter 5 are used to obtain an asymptotic expansion of the pdf of the sample latent roots of $B(B+W)^{-1}$ for large n_2 and $\Sigma^{-1}\Delta = n_2\Theta$. The same approach as in Chapter 4 is used to investigate the tests that the residual roots of Θ are zero.

1.6. Notation.

1. R^k stands for k-dimensional Euclidean space.

2. Lower case letters are used for scalars and column vectors. If x is an m-dimensional vector then x_i $(1 \leq i \leq m)$ are the components of x.

3. $O(m)$ is the group of all $m \times m$ orthogonal matrices. If $H \in O(m)$ then (dH) is the invariant Haar measure on $O(m)$ normalized so that $\int_{O(m)} (dH) = 1$.

4. $V(k,m)$ is the group of all $m \times k$ $(m \geq k)$ matrices with orthonormal columns. If $H_1 \in V(k,m)$ then (dH_1) is the invariant Haar measure on $V(k,m)$ normalized so that $\int_{V(k,m)} (dH_1) = 1$.

5. If X is an $m \times m$ matrix then $X > 0$ means X is positive definite.

6. If $X = (x_{ij})$ is an $m \times m$ matrix then (dX) stands for the exterior product of the elements of X. In particular, if X is symmetric then $(dX) = \Lambda_{i \leq j}^{m} dx_{ij}$ where Λ denotes the exterior product.

7. $\Gamma_p(a)$ is the multivariate gamma function defined by (1.2.4).

8. $\mathrm{diag}(A_1,\ldots,A_k)$ is a block diagonal matrix with $A_1,\ldots,A_k$ as diagonal elements. The A_i may be either scalars or square matrices.

9. O is used to denote a matrix of zeros. The dimension of O should be clear from the context.

10. I_m is the $m \times m$ identity matrix.

11. If $A = (a_{ij})$ is an $m \times m$ matrix, then

$$\mathrm{tr}(A) = \text{trace } A = \sum_{i=1}^{m} a_{ii} ,$$

$$\det A = \text{determinant of } A , \quad \text{and}$$

$$\mathrm{etr}(A) = \exp\{\mathrm{tr}(A)\} .$$

12. pdf stands for "probability density function".

CHAPTER 2

ASYMPTOTICS -- DEFINITIONS AND TECHNIQUES

2.1. Introduction.

One of the main objectives of this research is to derive asymptotic expansions for the hypergeometric functions which occur in the pdf's of the latent roots in canonical correlation analysis and in discriminant analysis (see Chapter 1, Sections 3 and 4). In this chapter we define "asymptotic expansion" as it applies to these hypergeometric functions and develop the techniques which will be used to derive the expansions.

2.2. Framework and Definitions.

Assume that

(i) A is a subset of R^p ;

(ii) N is the set of positive integers: and

(iii) $F(n,\alpha)$ is a real valued function defined on $N \times A$ $(n\in N, \alpha\in A)$.

In the specific problems considered in this research, F is a hypergeometric function of two matrix arguments, A is the set of sample and population latent roots, and n is the sample size minus some constant.

The problem is to approximate $F(n,\alpha)$ for large n . The following definitions have been adapted from Erdèlyi (1956).

Definition 2.2.1 --

The sequence of functions $\{\varphi_k(n,\alpha)\ ,\ k=1,2,\ldots\}$ is an asymptotic sequence as $n \to \infty$ if for each α

$$\varphi_{k+1}(n,\alpha) = o\{\varphi_k(n,\alpha)\} \qquad \text{as} \qquad n \to \infty .$$

Definition 2.2.2 --

Let $\{\varphi_k(n,\alpha)\ ,\ k=1,2,\ldots\}$ be an asymptotic sequence. The (formal)

series $\Sigma_k a_k\varphi_k(n,\alpha)$ is an asymptotic expansion to K terms of $F(n,\alpha)$ as $n \to \infty$ if for each α

$$F(n,\alpha) = \sum_{k=1}^{K} a_k\varphi_k(n,\alpha) + o\{\varphi_K(n,\alpha)\} \qquad \text{as} \qquad n \to \infty .$$

This is written as

$$F(n,\alpha) \sim \sum_{k=1}^{K} a_k\varphi_k(n,\alpha) .$$

Definition 2.2.3 --

The function $\varphi(n,\alpha)$ is an asymptotic representation for $F(n,\alpha)$ as $n \to \infty$ if for each α

$$F(n,\alpha) = \varphi(n,\alpha) + o\{\varphi(n,\alpha)\} \qquad \text{as} \qquad n \to \infty .$$

This is written as $F(n,\alpha) \sim \varphi(n,\alpha)$.

For the particular functions $F(n,\alpha)$ that we will consider, an asymptotic expansion will be obtained in two steps. First an asymptotic representation will be derived using an extension of Laplace's method for obtaining the asymptotic behavior of integrals. This asymptotic representation will then be used to derive partial differential equations satisfied by further terms in an asymptotic expansion of $F(n,\alpha)$.

2.3. Derivation of an Asymptotic Representation.

Assume that for each fixed $\alpha \in A$, $F(n,\alpha)$ has an integral representation of the form

$$F(n,\alpha) = \int_D \tau(x)\{f(x,\alpha)\}^n dx . \tag{2.3.1}$$

Hsu (1948) found an asymptotic representation for such integrals when D is a compact subset of R^m and τ and f satisfy certain regularity conditions. It is possible to show that the functions $F(n,\alpha)$ which we

will consider are asymptotically equivalent to an integral which satisfies the conditions of Hsu's lemma and hence an asymptotic representation $\varphi(n, \alpha)$ of $F(n, \alpha)$ can be obtained by applying Hsu's lemma to this integral. An easier approach is to reformulate Hsu's lemma, basically by replacing the conditions on the domain of integration D with additional conditions on the function f. In fact we will show that on a subset B of A, $F(n, \alpha) \sim \varphi(n, \alpha)$ uniformly in α. This means that given $\epsilon > 0$, there exists an $n(\epsilon)$, such that $n > n(\epsilon)$ implies $|F(n, \alpha)/\varphi(n, \alpha) - 1| < \epsilon$ for all $\alpha \in B$. The result we need is given in Theorem 2.3.1. Corollary 2.3.1 is essentially a reformulation of Hsu's lemma.

The following notation is used in Theorem 2.3.1 and its corollary. $S(r, x)$ is the open sphere with center x and radius r, i.e., $S(r, x) = \{y : (y-x)'(y-x) < r^2\}$. The arguments following a partial derivative of f are the values of x and α at which the partial derivatives are to be evaluated. For example,

$$\frac{\partial f}{\partial x_i}(\xi, a) \quad \text{means} \quad \frac{\partial f}{\partial x_i}(x, \alpha)$$

evaluated at $x = \xi$ and $\alpha = a$.

The following well known result (see, for example, Bellman (1960)) is used in the proof of Theorem 2.3.1.

<u>Lemma 2.3.1</u> --

Let Γ be an $m \times m$ real symmetric matrix with latent roots $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_m$. Then

$$\gamma_1 = \max_{x \neq 0} \frac{x'\Gamma x}{x'x} = \max_{x'x=1} x'\Gamma x$$

and

$$\gamma_m = \min_{x \neq 0} \frac{x'\Gamma x}{x'x} = \min_{x'x=1} x'\Gamma x .$$

Theorem 2.3.1 --

Let D be a subset of R^m and A a subset of R^p. Assume that $B \subset A$ and that τ and f are real valued functions defined on D and $D \times A$, respectively, such that

(i) there exists an integer $k \geq 0$ and a constant r_k such that $\tau(x)\{f(x,\alpha)\}^k$ is absolutely integrable on D and

$$\int_D |\tau(x)\{f(x,\alpha)\}^k| \, dx \leq r_k$$

for all $\alpha \in B$;

(ii) for each $\alpha \in B$, $f(x,\alpha)$ has an absolute maximum value at an interior point $\xi(\alpha)$ of D and

$$0 < \chi^{-1} = \inf_{\alpha \in B} f(\xi(\alpha),\alpha) ;$$

(iii) there exists a $\delta_1 > 0$ such that for every $\alpha \in B$, $S\{\delta_1, \xi(\alpha)\}$ is contained in the interior of D and $x \in S\{\delta_1, \xi(\alpha)\}$ implies $f(x,\alpha) > 0$;

(iv) there exists a $\delta_2 > 0$ such that for every $\alpha \in B$ all partial derivatives

$$\frac{\partial f}{\partial x_i}(x,\alpha) \quad \text{and} \quad \frac{\partial^2 f}{\partial x_i \partial x_j}(x,\alpha)$$

$(i,j=1,\ldots,m)$ exist and are continuous functions of x on $S\{\delta_2, \xi(\alpha)\}$;

(v) $0 < \mu^2 = \inf_{\alpha \in B} \gamma_m^{\,2}(\alpha)$ and $\sup_{\alpha \in B} \gamma_1^{\,2}(\alpha) = \kappa^2 < \infty$

where $\gamma_1^{\,2}(\alpha) \geq \cdots \geq \gamma_m^{\,2}(\alpha)$ are the latent roots of the $m \times m$ symmetric matric $\Omega\{\xi(\alpha),\alpha\}$. $\Omega(x,\alpha) = (\omega_{ij}(x,\alpha))$ is defined for all $\alpha \in B$ and $x \in S\{\delta_1, \xi(\alpha)\} \cap S\{\delta_2, \xi(\alpha)\}$ by

$$\omega_{ij}(x,\alpha) = -\frac{\partial^2 \ln f}{\partial x_i \partial x_j}(x,\alpha) ;$$

(vi) for every $\epsilon > 0$ there exists a $\delta_3 > 0$ such that for all $\alpha \in B$ and i, j $(1 \le i, j \le m)$, $x \in S\{\delta_3, \xi(\alpha)\}$ implies $|\omega_{ij}(x,\alpha) - \omega_{ij}\{\xi(\alpha),\alpha\}| < \epsilon$;

(vii) for every $\delta_4 > 0$ there exists a constant θ, $0 < \theta < 1$, such that $|f(x,\alpha)/f\{\xi(\alpha),\alpha\}| < \theta$ for all $x \in D - S\{\delta_4, \xi(\alpha)\}$ and $\alpha \in B$;

(viii) $0 < \inf\limits_{\alpha \in B} \tau\{\xi(\alpha)\}$; and

(ix) there exists a $\delta_5 > 0$ such that $\tau(x)$ is continuous on $\bigcup\limits_{\alpha \in B} S\{\delta_5, \xi(\alpha)\}$ and for every $\epsilon > 0$ there exists a $\delta_6 > 0$ such that $x \in S\{\delta_6, \xi(\alpha)\}$ implies $|\tau(x) - \tau\{\xi(\alpha)\}| < \epsilon$ for every $\alpha \in B$.

Then for large n

$$\int_D \tau(x)\,\{f(x,\alpha)\}^n dx \sim (2\pi/n)^{\frac{1}{2}m}[f\{\xi(\alpha),\alpha\}]^n \tau\{\xi(\alpha)\}[\Delta\{\xi(\alpha)\}]^{-\frac{1}{2}}$$

uniformly in $\alpha \in B$, where Δ denotes the Hessian of $-\ln f$ as a function of x, i.e.,

$$\Delta\{\xi(\alpha)\} = \det\left(-\frac{\partial^2 \ln f}{\partial x_i \partial x_j}\{\xi(\alpha),\alpha\}\right)$$

and "$a \sim b$ for large n" means $\lim\limits_{n \to \infty} \frac{a}{b} = 1$.

Proof --

The proof is based on Hsu's (1948) proof.

By conditions (ii) and (iv)

$$\frac{\partial f}{\partial x_i}\{\xi(\alpha),\alpha\} = 0 \qquad (i=1,\ldots,m)$$

and the $m \times m$ symmetric matric

$$\left(-\frac{\partial^2 f}{\partial x_i \partial x_j}\{\xi(\alpha),\alpha\}\right)$$

is positive definite for every $\alpha \in B$. By condition (iii) there exists

a $\delta_1 > 0$ such that for all $\alpha \in B$, $S\{\delta_1, \xi(\alpha)\}$ is contained in the interior of D and $x \in S\{\delta_1, \xi(\alpha)\}$ implies $f(x,\alpha) > 0$. Hence for each α we can define the real valued function $\Psi(x,\alpha)$ on $S\{\delta_1, \xi(\alpha)\}$ by $\Psi(x,\alpha) = \ln f(x,\alpha)$. By conditions (ii) and (iv) there exists a δ_2, $0 < \delta_2 \le \delta_1$, such that $\Psi(x,\alpha)$ is twice differentiable as a function of x on $S\{\delta_2, \xi(\alpha)\}$. In particular

$$\frac{\partial \Psi}{\partial x_i}\{\xi(\alpha),\alpha\} = [f\{\xi(\alpha),\alpha\}]^{-1} \frac{\partial f}{\partial x_i}\{\xi(\alpha),\alpha\} = 0 \qquad (i=1,\ldots,m)$$

and

$$\Omega\{\xi(\alpha),\alpha\} = \left(-\frac{\partial^2 \Psi}{\partial x_i \partial x_j}\{\xi(\alpha),\alpha\}\right) = [f\{\xi(\alpha),\alpha\}]^{-1}\left(-\frac{\partial^2 f}{\partial x_i \partial x_j}\{\xi(\alpha),\alpha\}\right)$$

is positive definite. We can expand $\Psi(x,\alpha)$ for $x \in S\{\delta_2, \xi(\alpha)\}$ in a Taylor series about $\xi(\alpha)$. The result is

$$(2.3.2) \qquad \Psi(x,\alpha) - \Psi\{\xi(\alpha),\alpha\} = -\tfrac{1}{2}\{x-\xi(\alpha)\}'\Omega\{\eta(\alpha),\alpha\}\{x-\xi(\alpha)\}$$

where $\eta(\alpha) = \xi(\alpha) + \beta(\alpha)\{x - \xi(\alpha)\}$ for some $\beta(\alpha)$, $0 \le \beta(\alpha) \le 1$.

First consider the integral

$$I_1(\alpha) = \int_D \left[\frac{f(x,\alpha)}{f\{\xi(\alpha),\alpha\}}\right]^n dx \qquad \text{where} \qquad \alpha \in B .$$

Let $G(\alpha)$ be any open subset of $S\{\delta_2, \xi(\alpha)\}$ containing $\xi(\alpha)$. Then from (2.3.2)

$$(2.3.3) \qquad I_1(\alpha) = \int_{G(\alpha)} \exp[-\tfrac{1}{2}n\{x-\xi(\alpha)\}'\Omega\{\eta(\alpha),\alpha\}\{x-\xi(\alpha)\}]dx$$

$$+ \int_{D-G(\alpha)} \left[\frac{f(x,\alpha)}{f\{\xi(\alpha),\alpha\}}\right]^n dx .$$

$\Omega\{\xi(\alpha),\alpha\}$ is positive definite and from condition (iv) $\omega_{ij}(x,\alpha)$ $(i,j=1,\ldots,m)$ is a continuous function of x on $S\{\delta_2, \xi(\alpha)\}$. Hence

$$1 = \lim_{x \to \xi(\alpha)} \frac{\{x-\xi(\alpha)\}'\Omega\{\eta(\alpha),\alpha\}\{x-\xi(\alpha)\}}{\{x-\xi(\alpha)\}'\Omega\{\xi(\alpha),\alpha\}\{x-\xi(\alpha)\}}$$

$$= \lim_{y \to 0} \frac{y'\Omega\{\xi(\alpha)+\beta(\alpha)y,\alpha\}y}{y'\Omega\{\xi(\alpha),\alpha\}y}$$

where $y = x - \xi(\alpha)$. This result means that

(2.3.4) $$y'\Omega\{\xi(\alpha)+\beta(\alpha)y,\alpha\}y = y'\Omega\{\xi(\alpha),\alpha\}y\{1+\lambda(y,\alpha)\}$$

where, for each fixed $\alpha \in B$, $\lambda(y,\alpha) \to 0$ as $y \to 0$. We want to show that $\lambda(y,\alpha) \to 0$ as $y \to 0$, uniformly in $\alpha \in B$. That is, given $\epsilon_0 > 0$ there exists a $\delta_0 > 0$ such that $y \in S(\delta_0, 0)$ implies $|\lambda(y,\alpha)| < \epsilon_0$ for all $\alpha \in B$. Suppose $\epsilon_0 > 0$ is given. By lemma 2.3.1 and condition (v)

$$y'\Omega\{\xi(\alpha),\alpha\}y \geq \gamma_m^2(\alpha)y'y$$

$$\geq \mu^2 y'y .$$

By condition (vi) there exists a $\delta_0 > 0$ such that $y \in S(\delta_0, 0)$ implies $|\omega_{ij}\{\xi(\alpha)+\beta(\alpha)y,\alpha\} - \omega_{ij}\{\xi(\alpha),\alpha\}| < \mu^2\epsilon_0/m^2$ for all i,j $(1 \leq i,j \leq m)$ and for all $\alpha \in B$. Therefore for all nonzero $y \in S(\delta_0, 0)$ and $\alpha \in B$

$$|\lambda(y,\alpha)| = \frac{|y'[\Omega\{\xi(\alpha)+\beta(\alpha)y,\alpha\} - \Omega\{\xi(\alpha),\alpha\}]y|}{y'\Omega\{\xi(\alpha),\alpha\}y}$$

$$\leq \frac{\sum_{i=1}^{m}\sum_{j=1}^{m} |y_i||y_j|\mu^2\epsilon_0/m^2}{\mu^2 y'y}$$

$$\leq \frac{m^2\epsilon_0(\max|y_i|)^2}{m^2 y'y}$$

$$\leq \frac{\epsilon_0 y'y}{y'y} = \epsilon_0$$

That is, $\lambda(y,\alpha) \to 0$ as $y \to 0$ uniformly in $\alpha \in B$.

Make the transformation of variables $y = x - \xi(\alpha)$ in (2.3.3). Then

$$I_1(\alpha) = \int_{F(\alpha)} \{g(y,\alpha)\}^n dy + \int_{E(\alpha)-F(\alpha)} \{h(y,\alpha)\}^n dy \tag{2.3.5}$$

where
$$g(y,\alpha) = \exp[-\tfrac{1}{2}ny'\Omega\{\xi(\alpha),\alpha\}y\ \{1+\lambda(y,\alpha)\}] ,$$

$$h(y,\alpha) = \frac{f\{\xi(\alpha)+y,\alpha\}}{f\{\xi(\alpha),\alpha\}} ,$$

$F(\alpha) = \{x-\xi(\alpha) : x \in G(\alpha)\}$ and $E(\alpha) = \{x-\xi(\alpha) : x \in D\}$.

Since $\Omega\{\xi(\alpha),\alpha\}$ is positive definite there exists an $m \times m$ orthogonal matrix $H(\alpha)$ such that $H(\alpha)'\Omega\{\xi(\alpha),\alpha\}H(\alpha) = \Gamma^2(\alpha)$, where $\Gamma(\alpha) = \mathrm{diag}\{\gamma_1(\alpha),\ldots,\gamma_m(\alpha)\}$ with $\gamma_1(\alpha) \geq \gamma_2(\alpha) \geq \cdots \geq \gamma_m(\alpha) > 0$. Let

$$M(\alpha) = H(\alpha)\Gamma(\alpha)^{-1}$$

and

$$z = M(\alpha)^{-1}\ y .$$

Then

$$y'\Omega\{\xi(\alpha),\alpha\}y = z'M(\alpha)'\Omega\{\xi(\alpha),\alpha\}M(\alpha)z = z'z .$$

The Jacobian of this transformation $J(y \to z)$ is

$$J(y \to z) = \det M(\alpha) = \det \Omega\{\xi(\alpha),\alpha\}^{-\frac{1}{2}} = \Delta\{\xi(\alpha)\}^{-\frac{1}{2}} .$$

Let $\epsilon > 0$ be given. By lemma 2.3.1 and condition (v)

$$\mu^2 y'y \leq y'\Omega\{\xi(\alpha),\alpha\}y = z'z \leq \kappa^2 y'y \tag{2.3.6}$$

for all $\alpha \in B$. Since $\lambda(y,\alpha) \to 0$ as $y \to 0$ uniformly in $\alpha \in B$ there exists a δ , $0 < \delta \leq \delta_2$, such that $y \in S(\delta,0)$ implies $|\lambda(y,\alpha)| < \epsilon$ for all $\alpha \in B$. Let

$$D(\delta) = \{z : |z_i| < \delta\ \mu/m^{\frac{1}{2}},\ i=1,2,\ldots,m\}$$

and let $C(\delta,\alpha)$ be the inverse image of $D(\delta)$ under the transformation $z = M(\alpha)^{-1}y$. Since $D(\delta)$ is open and the transformation is continuous, $C(\delta,\alpha)$ is open. If $z \in D(\delta)$ then $z'z < \delta^2\mu^2$ and so by (2.3.6), if

$y \in C(\delta,\alpha)$ then $y'y < \delta^2$. That is, $C(\delta,\alpha) \subset S(\delta,0)$. If $y \in S(\delta\,\mu/(\kappa m^{\frac{1}{2}}),0)$ then $y'y < \delta^2\mu^2/(\kappa^2 m)$ and hence by (2.3.6), $z'z < \delta^2\mu^2/m$, i.e., $|z_i| < \delta\,\mu/m^{\frac{1}{2}}$, $i=1,\ldots,m$. Therefore $S(\delta\,\mu/(\kappa m^{\frac{1}{2}}),0) \subset C(\delta,\alpha)$. In summary

$$S(\delta\,\mu/(\kappa m^{\frac{1}{2}}),0) \subset C(\delta,\alpha) \subset S(\delta,0)\,. \tag{2.3.7}$$

Let $\eta(z,\alpha)$ be the function defined by $\lambda(y,\alpha) = \eta(z,\alpha)$ where $z = M(\alpha)^{-1}y$. If $z \in D(\delta)$ then $y \in C(\delta,\alpha) \subset S(\delta,0)$ and therefore $|\eta(z,\alpha)| = |\lambda(y,\alpha)| < \epsilon$ for all $\alpha \in B$. That is, $\eta(z,\alpha) \to 0$ as $z \to 0$ uniformly in $\alpha \in B$.

Since $\delta \le \delta_2$ it follows from (2.3.5) with $F(\alpha) = C(\delta,\alpha)$ that

$$I_1(\alpha) = \int_{C(\delta,\alpha)} \{g(y,\alpha)\}^n dy + \int_{E(\alpha)-C(\delta,\alpha)} \{h(y,\alpha)\}^n\, dy\,. \tag{2.3.8}$$

Let

$$I_2(\alpha) = \int_{C(\delta,\alpha)} \{g(y,\alpha)\}^n\, dy$$

and make the transformation of variables $z = M(\alpha)^{-1}y$ in $I_2(\alpha)$. Then

$$\begin{aligned} I_2(\alpha) &= \Delta\{\xi(\alpha)\}^{-\frac{1}{2}} \int_{D(\delta)} [g(M(\alpha)z,\alpha)]^n dz \\ &= \Delta\{\xi(\alpha)\}^{-\frac{1}{2}} \int_{-d}^{d} \cdots \int_{-d}^{d} \exp[-\tfrac{1}{2}nz'z\{1+\eta(z,\alpha)\}]dz \end{aligned}$$

where $d = \delta\,\mu/m^{\frac{1}{2}}$. δ was chosen so that $z \in D(\delta)$ implies $|\eta(z,\alpha)| < \epsilon$ for all $\alpha \in B$. Therefore

$$\begin{aligned} \Delta\{\xi(\alpha)\}^{-\frac{1}{2}} \prod_{k=1}^{m} \int_{-d}^{d} \exp[-\tfrac{1}{2}nz_k^{\,2}(1+\epsilon)]dz_k &\le I_2(\alpha) \\ &\le \Delta\,\{\xi(\alpha)\}^{-\frac{1}{2}} \prod_{k=1}^{m} \int_{-d}^{d} \exp[-\tfrac{1}{2}nz_k^{\,2}(1-\epsilon)]dz_k\,. \end{aligned}$$

The integrals in the above inequalities are, up to a multiplicative constant, integrals of normal density functions. Hence

$$(2.3.9) \qquad \left[\frac{2\pi(1-a_n)^2}{n(1+\epsilon)}\right]^{\frac{1}{2}m} \Delta\{\xi(\alpha)\}^{-\frac{1}{2}} \le I_2(\alpha) \le \left[\frac{2\pi(1-b_n)^2}{n(1-\epsilon)}\right]^{\frac{1}{2}m} \Delta\{\xi(\alpha)\}^{-\frac{1}{2}}$$

for all $\alpha \in B$, where

$$a_n = 2\Phi\{-\delta\, \mu n^{\frac{1}{2}}(1+\epsilon)^{\frac{1}{2}}/m^{\frac{1}{2}}\} ,$$

$$b_n = 2\Phi\{-\delta\, \mu n^{\frac{1}{2}}(1-\epsilon)^{\frac{1}{2}}/m^{\frac{1}{2}}\} ,$$

and

$$\Phi(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt .$$

Let $\epsilon > 0$ be defined as before and consider the integral

$$J_1(\alpha) = \int_D \tau(x) \left[\frac{f(x,\alpha)}{f\{\xi(\alpha),\alpha\}}\right]^n dx .$$

From the previous argument it follows that

$$(2.3.10) \quad J_1(\alpha) = \int_{C(\delta,\alpha)} \tau\{\xi(\alpha)+y\}\{g(y,\alpha)\}^n dy + \int_{E(\alpha)-C(\delta,\alpha)} \tau\{\xi(\alpha)+y\}\{h(y,\alpha)\}^n dy .$$

Let

$$J_2(\alpha) = \int_{C(\delta,\alpha)} \tau\{\xi(\alpha)+y\}\{g(y,\alpha)\}^n dy .$$

By condition (ix) and the first mean value theorem we have

$$(2.3.11) \qquad J_2(\alpha) = \tau\{\xi(\alpha)+\zeta(\alpha)\} \int_{C(\delta,\alpha)} \{g(y,\alpha)\}^n dy$$

where $\zeta(\alpha) \in C(\delta,\alpha)$. Both (2.3.10) and (2.3.11) hold for arbitrary δ, $0 < \delta < \min(\delta_2, \delta_5)$. By condition (ix) we can choose δ small enough so that $\zeta(\alpha) \in C(\delta,\alpha)$ implies $|\tau\{\xi(\alpha)+\zeta(\alpha)\}-\tau\{\xi(\alpha)\}| < \epsilon$ for all $\alpha \in B$. Then by (2.3.9)

$$(2.3.12) \quad \frac{[\tau\{\xi(\alpha)\}-\epsilon]}{\Delta\{\xi(\alpha)\}^{\frac{1}{2}}} \left[\frac{2\pi(1-a_n)^2}{n(1+\epsilon)}\right]^{\frac{1}{2}m} \le J_2(\alpha) \le \frac{[\tau\{\xi(\alpha)\}+\epsilon]}{\Delta\{\xi(\alpha)\}^{\frac{1}{2}}} \left[\frac{2\pi(1-b_n)^2}{n(1-\epsilon)}\right]^{\frac{1}{2}m} .$$

By condition (i) $\tau\{\xi(\alpha)+y\}[f\{\xi(\alpha)+y,\alpha\}]^k$ is absolutely integrable over $E(\alpha)$ and

$$\int_{E(\alpha)} |\tau\{\xi(\alpha)+y\}[f\{\xi(\alpha)+y,\alpha\}]^k| dy \le r_k .$$

Let $S = S\{\delta\,\mu/(\kappa m^{\frac{1}{2}}), 0\}$. Then $\tau\{\xi(\alpha)+y\}[f\{\xi(\alpha)+y,\alpha\}]^k$ is absolutely integrable over $E(\alpha) - S$ and

$$\int_{E(\alpha)-S} |\tau\{\xi(\alpha)+y\}[f\{\xi(\alpha)+y,\alpha\}]^k| dy \le r_k .$$

From (2.3.7) $S \subset C(\delta,\alpha)$ and by condition (vii) there exists a θ, $0 < \theta < 1$, such that

$$|h(y,\alpha)| < \theta$$

for all $\alpha \in B$ and $y \in E(\alpha) - S$.

Then for $n \ge k$ and all $\alpha \in B$

$$(2.3.13) \qquad \Big|\int_{E(\alpha)-C(\delta,\alpha)} \tau\{\xi(\alpha)+y\}\{h(y,\alpha)\}^n dy\Big|$$

$$\le \int_{E(\alpha)-S} |\tau\{\xi(\alpha)+y\}|\;|h(y,\alpha)|^n dy$$

$$\le \int_{E(\alpha)-S} |\tau\{\xi(\alpha)+y\}|\,|f\{\xi(\alpha)+y,\alpha\}|^k \,\frac{\theta^{n-k}}{f\{\xi(\alpha),\alpha\}^k}\, dy$$

$$\le r_k\,\theta^{n-k}\chi^k$$

where the last line follows from condition (ii). Combining (2.3.13) and (2.3.12) we have

$$(2.3.14) \quad \frac{[\tau\{\xi(\alpha)\}-\epsilon]}{\tau\{\xi(\alpha)\}}\left[\frac{-t_n\Delta\{\xi(\alpha)\}^{\frac{1}{2}}}{\tau\{\xi(\alpha)\}-\epsilon} + \left\{\frac{(1-a_n)^2}{1+\epsilon}\right\}^{\frac{1}{2}m}\right] \le \frac{J_1(\alpha)\Delta\{\xi(\alpha)\}^{\frac{1}{2}}}{\tau\{\xi(\alpha)\}}\,(n/2\pi)^{\frac{1}{2}m}$$

$$\le \frac{[\tau\{\xi(\alpha)\}+\epsilon]}{\tau\{\xi(\alpha)\}}\left[\frac{t_n\Delta\{\xi(\alpha)\}^{\frac{1}{2}}}{\tau\{\xi(\alpha)\}+\epsilon} + \left\{\frac{(1-b_n)^2}{1-\epsilon}\right\}^{\frac{1}{2}m}\right]$$

for all $\alpha \in B$, where $t_n = r_k \theta^{n-k} \chi^k (n/2\pi)^{\frac{1}{2}m}$. Note that a_n, b_n, and t_n do not depend on α and $a_n \to 0$, $b_n \to 0$, and $t_n \to 0$ as $n \to \infty$. Since $\varepsilon > 0$ was arbitrary it follows from (2.3.14) and conditions (v) and (viii) that

$$J_1(\alpha) \sim \tau\{\xi(\alpha)\}\ \Delta\{\xi(\alpha)\}^{-\frac{1}{2}}(2\pi/n)^{\frac{1}{2}m}$$

uniformly in $\alpha \in B$, proving the theorem.

Corollary 2.3.1 --

Let D be a subset of R^m and let τ and f be real valued functions defined on D such that

(i) there exists a $k \geq 0$ such that $\tau(x)\{f(x)\}^k$ is absolutely integrable on D;

(ii) $f(x)$ has an absolute *maximum* value at an interior point ξ of D and $f(\xi) > 0$;

(iii) all partial derivatives

$$\frac{\partial f}{\partial x_i} \quad \text{and} \quad \frac{\partial^2 f}{\partial x_i \partial x_j} \qquad (i,j=1,\ldots,m)$$

exist and are continuous in a neighborhood of ξ;

(iv) for every neighborhood N of ξ there exists a constant θ, $0 < \theta < 1$, such that $|f(x)/f(\xi)| < \theta$ for all $x \in D - N$; and

(v) τ is continuous in a neighborhood of ξ and $\tau(\xi) \neq 0$.

Then for large n

$$\int_D \tau(x)\ \{f(x)\}^n dx \sim (2\pi/n)^{\frac{1}{2}m}\{f(\xi)\}^n \tau(\xi)[\Delta(\xi)]^{-\frac{1}{2}}.$$

Proof--

The corollary follows from Theorem 2.3.1 by taking B to be a set consisting of 1 point.

2.4. Derivation of an Asymptotic Expansion.

In this section $F(n,\alpha)$ will denote one of the hypergeometric functions which occur in the pdf of the latent roots in canonical correlation analysis and in discriminant analysis.

In principle, further terms in an asymptotic expansion of $F(n,\alpha)$ could be obtained by a more detailed analysis of the integral (2.3.1). In practice the mathematics is intractable, but it is easy to see that such an analysis would lead to an asymptotic expansion of the form $F(n,\alpha) \sim \varphi(n,\alpha)G(n,\alpha)$ where

$$G(n,\alpha) = 1 + \frac{P_1(\alpha)}{n} + \frac{P_2(\alpha)}{n^2} + \cdots \tag{2.4.1}$$

and $\varphi(n,\alpha)$ is the asymptotic representation of $F(n,\alpha)$ obtained from Theorem 2.3.1.

By extending the results of Constantine and Muirhead (1972) we obtain a differential equation satisfied by $F(n,\alpha)$ from which we derive a differential equation satisfied by $G(n,\alpha)$. The equation for $G(n,\alpha)$ turns out to have the form

$$(n\Delta_1 + \Delta_2)G(n,\alpha) = h(\alpha)G(n,\alpha) \tag{2.4.2}$$

where Δ_1 and Δ_2 are differential operators in the elements of α which do not depend on n. Substituting (2.4.1) into (2.4.2) and equating coefficients of like powers of n^{-1} gives a recursive system of differential equations for the P_i :

$$\begin{aligned} &\Delta_1 P_1 = h \\ &\Delta_1 P_i + (\Delta_2 - h)P_{i-1} = 0 \qquad i=2,3,\ldots \end{aligned} \tag{2.4.3}$$

The system of equations is solved recursively for the P_i . This procedure is justified by the uniqueness of asymptotic power series, i.e. if

$$0 \sim \sum_{i=1}^{\infty} \{\Delta_1 P_i + (\Delta_2 - h) P_{i-1}\} n^{-i}$$

then $\Delta_1 P_i + (\Delta_2 - h) P_{i-1} \equiv 0$.

It turns out that Δ_1 is a first order linear differential operator. It is well known from the theory of first order linear partial differential equations that the general solution of the equation

$$\sum_{i=1}^{p} f_i \frac{\partial P}{\partial \alpha_i} = R \,, \tag{2.4.4}$$

when the f_i are independent of P , is

$$P = Q + \psi(u_1, \ldots, u_{p-1}) \tag{2.4.5}$$

where Q is any particular solution of (2.4.4) and ψ is an abritrary function, subject to certain regularity conditions (see Goursat (1917)). The u_i are any $p - 1$ independent solutions of the system of ordinary differential equations given by

$$\frac{d\alpha_1}{f_1} = \frac{d\alpha_2}{f_2} = \cdots = \frac{d\alpha_p}{f_p} \,. \tag{2.4.6}$$

CHAPTER 3

AN ASYMPTOTIC EXPANSION OF ${}_2F_1(\frac{1}{2}n,\frac{1}{2}n;\frac{1}{2}q;P^2,R^2)$ FOR LARGE n.

3.1. Introduction.

Let $R = \mathrm{diag}(r_1,\ldots,r_p)$ be the diagonal matrix of the sample canonical correlation coefficients calculated from a sample of size $N = n + 1$ from a $(p+q)$-variate normal population $(n \geq p+q,\ q \geq p)$ and let $P = \mathrm{diag}(\rho_1,\ldots,\rho_p)$ be the diagonal matrix of the population correlation coefficients. The joint pdf of $r_1^2,\ldots,r_p^2$ involves the hypergeometric function ${}_2F_1(\frac{1}{2}n,\frac{1}{2}n;\frac{1}{2}q;P^2,R^2)$ which will be denoted in this chapter as ${}_2F_1(n,P^2,R^2)$. Assume that $1 > r_1 > \cdots > r_p > 0$, which is true with probability 1, and that $P = \mathrm{diag}(P_1,0)$ where $P_1 = \mathrm{diag}(\rho_1,\ldots,\rho_k)$ $(0 \leq k \leq p)$ with $1 > \rho_1 > \cdots > \rho_k > 0$.

The technique presented in Chapter 2 is used to derive an asymptotic expansion, up to and including terms of order n^{-1}, for ${}_2F_1(n,P^2,R^2)$. In addition the partial differential equations, derived by Constantine and Muirhead (1972), for the ${}_2F_1$, ${}_1F_1$, ${}_0F_1$, ${}_1F_0$, and ${}_0F_0$ hypergeometric functions of two symmetric matrices are generalized to include the case where one of the matrices P has the form $P = \mathrm{diag}(P_1,0)$.

3.2. An Asymptotic Representation of ${}_2F_1(\frac{1}{2}n,\frac{1}{2}n;\frac{1}{2}q;P^2,R^2)$ for Large n.

An asymptotic representation of ${}_2F_1(n,P^2,R^2)$ follows from Theorem 2.3.1 once a suitable integral representation has been found. Deriving such an integral representation and checking the conditions of Theorem 2.3.1 turns out to be a rather lengthy process.

We begin by using the results of Chapter 1, Section 2, to express ${}_2F_1(n,P^2,R^2)$ as a multiple integral. It follows from (1.2.6) that

$$ {}_2F_1(n,P^2,R^2) = \int_{O(p)} {}_2F_1^{(p)}(\tfrac{1}{2}n,\tfrac{1}{2}n;\tfrac{1}{2}q;PH'R^2HP)(dH) \ . $$

Partition H into two submatrices H_1 and H_2 consisting of the first k and last $p-k$ columns, respectively. It follows from (1.2.5) and the fact that $P = \mathrm{diag}(P_1, 0)$ that

(3.2.1) $$ {}_2F_1(n,P^2,R^2) = \int_{O(p)} {}_2F_1^{(k)}(\tfrac{1}{2}n,\tfrac{1}{2}n;\tfrac{1}{2}q;P_1H_1'R^2H_1P_1)(dH) \ . $$

Because the integrand does not depend on H_2 , we can integrate over H_2 using Lemma 3.2.1 which is due to Constantine and Muirhead (1976). Eventually ${}_2F_1(n,P^2,R^2)$ will be expressed as an integral of the form

$$ \int_D \tau(x)\{f(x,\rho,r)\}^n dx $$

where $\rho = (\rho_1,\ldots,\rho_k)'$ and $r = (r_1,\ldots,r_p)'$. Theorem 2.3.1 requires that the maximum of $f(x,\rho,r)$, for fixed ρ and r , be obtained at a unique interior point of D . It turns out that at the maximum H_1 is uniquely determined, but H is only determined up to a subspace defined by the orthogonal complement of H_1 .

Lemma 3.2.1--

Suppose

(i) Q_1 is the submatrix consisting of the first k columns of $Q \in O(p)$;

(ii) (dQ_1) is the invariant measure on $V(k,p)$;

(iii) $G = G(Q_1)$ is any $p \times (p-k)$ matrix with orthonormal columns orthogonal to Q_1 ;

(iv) $f(Q)$ is a function of Q ; and

(v) $H \in O(p-k)$.

Then

$$\int_{O(p)} f(Q)(dQ) = \int_{V(k,p)} \int_{O(p-k)} f(Q_1, GH)(dH)(dQ_1) .$$

In particular if $f(Q) = f(Q_1)$ then

$$\int_{O(p)} f(Q_1)(dQ) = \int_{V(k,p)} f(Q_1)(dQ_1) .$$

Applying Lemma 3.2.1 to (3.2.1) we have

$${}_2F_1(n, P^2, R^2) = \int_{V(k,p)} {}_2F_1^{(k)}(\tfrac{1}{2}n, \tfrac{1}{2}n; \tfrac{1}{2}q; P_1H_1'R^2H_1P_1)(dH_1) .$$

For $n > k - 1$ we may apply the Laplace transform relation (1.2.3) twice to the integrand in the previous integral to obtain

$${}_2F_1(n, P^2, R^2) = \{\Gamma_k(\tfrac{1}{2}n)\}^{-2} \int_{V(k,p)} \int_{X>0} \int_{Y>0} \operatorname{etr}(-X-Y) \det(XY)^{\frac{1}{2}(n-k-1)}$$

$$\times\ {}_0F_1(\tfrac{1}{2}q; Y^{\frac{1}{2}}X^{\frac{1}{2}}P_1H_1'R^2H_1P_1X^{\frac{1}{2}}Y^{\frac{1}{2}})(dY)(dX)(dH_1)$$

where X and Y are $k \times k$ positive definite matrices. By applying Bessel's integral (1.2.7) to the ${}_0F_1$ function we get

$${}_2F_1(n, P^2, R^2) = \{\Gamma_k(\tfrac{1}{2}n)\}^{-2} \int_{V(k,p)} \int_{X>0} \int_{Y>0} \int_{O(q)} \operatorname{etr}(-X-Y) \det(XY)^{\frac{1}{2}(n-k-1)}$$

$$\times \operatorname{etr}\{2[Y^{\frac{1}{2}}X^{\frac{1}{2}}P_1H_1'R:0]M_1\}(dM)(dY)(dX)(dH_1)$$

where M_1 is the $q \times k$ matrix formed by the first k columns of $M \in O(q)$ and 0 in $[Y^{\frac{1}{2}}X^{\frac{1}{2}}P_1H_1'R:0]$ is the $k \times (q-k)$ zero matrix. The integrand does not depend on the last $q - k$ columns of M and, for the same reason that we integrated over H_2 in (3.2.1), we integrate over these $q - k$ columns using Lemma 3.2.1. The result is

$$(3.2.2) \quad {}_2F_1(n, P^2, R^2) = \{\Gamma_k(\tfrac{1}{2}n)\}^{-2} \int_{V(k,p)} \int_{X>0} \int_{Y>0} \int_{V(k,q)} \operatorname{etr}(-X-Y)\det(XY)^{\frac{1}{2}(n-k-1)}$$

$$\times \operatorname{etr}\{2[Y^{\frac{1}{2}}X^{\frac{1}{2}}P_1H_1'R:0]M_1\}(dM_1)(dY)(dX)(dH_1) .$$

It would be difficult in the subsequent analysis to work directly with the positive definite matrices X and Y. However, X and Y have distinct latent roots except on sets of measure zero with respect to the measures (dX) and (dY), respectively. Therefore we can write

$$X = \tfrac{1}{2}nG'V^2G \qquad \text{and} \qquad Y = \tfrac{1}{2}nQ'U^2Q \tag{3.2.3}$$

where G, $Q \in O(k)$, $V = \mathrm{diag}(v_1, v_2, \ldots, v_k)$ with $v_1 > v_2 > \cdots > v_k > 0$, and $U = \mathrm{diag}(u_1, u_2, \ldots, u_k)$ with $u_1 > u_2 > \cdots > u_k > 0$. The transformation between X and (V,G) defined by (3.2.3) is one to one if G is restricted to orthogonal matrices with positive elements in the first column. Since the set of $G \in O(k)$ with zeros in the first column constitutes a set of zero (dG)-measure, we may assume, without loss of generality, that the elements in the first column of G are positive. By the same argument we may also assume that the elements in the first column of Q are positive and the transformation $Y \to (U,Q)$ is one to one. Let $J\{X \to (V,G)\}$ and $J\{Y \to (U,Q)\}$ be the Jacobians of these transformations. Then from Deemer and Olkin (1951)

$$J\{X \to (V,G)\} = \frac{(\tfrac{1}{2}n)^{\frac{1}{2}k(k+1)} 2^{2k} \pi^{\frac{1}{2}k^2}}{\Gamma_k(\tfrac{1}{2}k)} \det V \prod_{i<j}^{k} (v_i^2 - v_j^2)$$

and

$$J\{Y \to (U,Q)\} = \frac{(\tfrac{1}{2}n)^{\frac{1}{2}k(k+1)} 2^{2k} \pi^{\frac{1}{2}k^2}}{\Gamma_k(\tfrac{1}{2}k)} \det U \prod_{i<j}^{k} (u_i^2 - u_j^2) .$$

When we make this change of variables in (3.2.2), the resulting integral will involve integration with respect to (dG) and (dQ) over the subset $O_1(k)$ of $O(k)$ consisting of orthogonal $k \times k$ matrices with

positive elements in the first column. The only part of the resulting integrand which depends on G or Q is λ where $\lambda = \mathrm{etr}\{n[Q'UQG'VGP_1H_1'R:0]M_1\}$. Consider transformations of the form $G \to S_1G$ and $Q \to S_2Q$ where S_1 and S_2 are diagonal matrices with ± 1 along the diagonal. If the i-th diagonal element of S_1 is -1 then the transformation $G \to S_1G$ changes the sign of the elements in the i-th row of G. The same is true of S_2 and the transformation $Q \to S_2Q$. λ is invariant under these transformations and therefore the integral over $O_1(k)$ with respect to (dG) and (dQ) can be replaced by 2^{-2k} times the integral over $O(k)$ with respect to (dG) and (dQ). It follows after some simplification that (3.2.2) may be expressed as

$$(3.2.4) \qquad {}_2F_1(n, P^2, R^2) = C_n \int_\Lambda \Pi(y)\{g(y,\rho,r)\}^n dy$$

where

$$C_n = \frac{(\tfrac{1}{2}n)^{nk}\pi^{k^2}2^{2k}}{\{\Gamma_k(\tfrac{1}{2}n)\Gamma_k(\tfrac{1}{2}k)\}^2},$$

$$\Pi(y) = \prod_{i<j}^{k}\{(v_i^2-v_j^2)(u_i^2-u_j^2)\}\det(UV)^{-k},$$

$$g(y,\rho,r) = \mathrm{etr}(-\tfrac{1}{2}U^2-\tfrac{1}{2}V^2+[Q'UQG'VGP_1H_1'R:0]M_1)\det(UV),$$

$$\Lambda = V(k,p)\times O(k)\times D_V\times O(k)\times D_U\times V(k,q),$$

$$D_X = \{(x_1,\ldots,x_k) : x_1 > x_2 > \cdots > x_k > 0\},$$

$$\rho = (\rho_1,\ldots,\rho_k)', \quad r = (r_1,\ldots,r_\rho)'$$

and y is a point in Λ.

Finally, consider the transformations $G \to G$, $QG' \to F$, $M_1Q' \to E_1$, and $H_1 \to H_1$. Then $G,F \in O(k)$, $H_1 \in V(k,p)$, $E_1 \in V(k,q)$ and

$(dM_1)(dQ)(dG)(dH_1) = (dE_1)(dF)(dG)(dH_1)$. Also, since $\text{tr}(AB) = \text{tr}(BA)$, we have

$$\begin{aligned}\lambda &= \text{etr}\{n[Q'UQG'VGP_1H_1{}'R:0]M_1\} \\ &= \text{etr}\{n[UQG'VGP_1H_1{}'R:0]M_1Q'\} \\ &= \text{etr}\{n[UFVGP_1H_1{}'R:0]E_1\} .\end{aligned}$$

Making this transformation of variables in (3.2.4) gives

$$(3.2.5) \qquad {}_2F_1(n, P^2, R^2) = C_n \int_\Lambda \Pi(y)\{h(y,\rho,r)\}^n dy$$

where

$$h(y,\rho,r) = \text{etr}(-\tfrac{1}{2}U^2 - \tfrac{1}{2}V^2 + [UFVGP_1H_1{}'R:0]E_1)\det(UV) .$$

We now determine the maximum value of h for fixed P and R and show that the maximum is obtained at a finite number of interior points of Λ . The result is contained in the following lemma.

Lemma 3.2.2--

Let

(i) $P_1 = \text{diag}(\rho_1,\ldots,\rho_k)$ with $\rho_1 > \rho_2 > \cdots > \rho_k > 0$;

(ii) $R = \text{diag}(r_1,\ldots,r_p)$ with $r_1 > r_2 > \cdots > r_p > 0$ $(p \geq k)$;

(iii) $U = \text{diag}(u_1,\ldots,u_k)$ with $u_1 > u_2 > \cdots > u_k > 0$;

(iv) $V = \text{diag}(v_1,\ldots,v_k)$ with $v_1 > v_2 > \cdots > v_k > 0$;

(v) $F = (f_{ij})$ and $G = (g_{ij})$ $(1 \leq i,j \leq k)$ with $F, G \in O(k)$;

(vi) $E_1 = (e_{ij})$ $(1 \leq i \leq q, 1 \leq j \leq k)$ with $E_1 \in V(k,q)$;

(vii) $H_1 = (h_{ij})$ $(1 \leq i \leq p, 1 \leq j \leq k)$ with $H_1 \in V(k,p)$; and

(viii) $h = \text{etr}(-\tfrac{1}{2}U^2 - \tfrac{1}{2}V^2 + [UFVGP_1H_1{}'R:0]E_1)\det(UV)$.

Then the maximum value of h for fixed P_1 and R is

$$e^{-k} \prod_{i=1}^{k} (1 - r_i\rho_i)^{-1}$$

and the maximum is obtained if and only if $u_i = v_i = (1-\rho_i r_i)^{-\frac{1}{2}}$ $(1 \leq i \leq k)$, $G = \mathrm{diag}(\pm 1, \ldots, \pm 1)$, $F = \mathrm{diag}(\pm 1, \ldots, \pm 1)$,

$$H_1 = \begin{bmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \\ \cdot & \cdot\ \cdot & \cdot \\ & 0 & \end{bmatrix} \quad \text{and} \quad E_1 = \begin{bmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \\ \cdot & \cdot\ \cdot & \cdot \\ & 0 & \end{bmatrix}$$

where G, F, H_1, and E_1 satisfy the following constraints

$$f_{ii} g_{ii} h_{ii} e_{ii} = 1 \qquad (1 \leq i \leq k) .$$

Proof--

Let $h_1 = e^T$, where $T = \mathrm{tr}([UFVGP_1H_1'R:0]E_1)$, and $h_2 = \mathrm{etr}(-\frac{1}{2}U^2-\frac{1}{2}V^2)\det(UV)$. Then $h = h_1h_2$. Maximizing T is equivalent to maximizing h_1 since h_1 is a strictly increasing function of T . It follows from Theorem 2 of the appendix, that for fixed U, V, P_1, and R , the maximum value of T is

$$\sum_{i=1}^{k} u_i v_i r_i \rho_i$$

and the maximum is obtained if and only if $F = \mathrm{diag}(\pm 1, \ldots, \pm 1)$, $G = \mathrm{diag}(\pm 1, \ldots, \pm 1)$,

$$H_1 = \begin{bmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \\ \cdot & \cdot\ \cdot & \cdot \\ & 0 & \end{bmatrix} \quad \text{and} \quad E_1 = \begin{bmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \\ \cdot & \cdot\ \cdot & \cdot \\ & 0 & \end{bmatrix}$$

where F, G, H_1 and E_1 satisfy the constraints $f_{ii} g_{ii} h_{ii} e_{ii} = 1$ $(1 \leq i \leq k)$. Note that the location of the maxima does not depend on U and V . Consequently, h can be maximized in two stages as follows:

$$\max_{U,V,F,G,H_1,E_1} h = \max_{U,V}\left(h_2 \left(\max_{F,G,H_1,E_1} h_1\right)\right) .$$

The problem of maximizing h thus reduces to maximizing

$$g_1(U,V) = \exp\{\sum_{i=1}^{k}(-\tfrac{1}{2}u_i^2 - \tfrac{1}{2}v_i^2 + u_i v_i r_i \rho_i)\} \prod_{i=1}^{k} u_i v_i$$

$$= \prod_{i=1}^{k}\{\exp(-\tfrac{1}{2}u_i^2 - \tfrac{1}{2}v_i^2 + u_i v_i r_i \rho_i + \ln u_i + \ln v_i)\} .$$

Consider the function $g_2(x,y)$ defined for all $x > 0$ and $y > 0$ by

$$g_2(x,y) = -\tfrac{1}{2}x^2 - \tfrac{1}{2}y^2 + xy\alpha + \ln x + \ln y$$

with $0 < \alpha < 1$. It is easily shown that g_2 has an absolute maximum value of $e^{-1}(1-\alpha)^{-1}$ which is obtained if and only if $x = y = (1-\alpha)^{-\frac{1}{2}}$. Lemma 3.2.2 follows from this result and the fact that $(1-r_1\rho_1)^{-\frac{1}{2}} > (1-r_2\rho_2)^{-\frac{1}{2}} > \cdots > (1-r_k\rho_k)^{-\frac{1}{2}} > 0$.

At each of the points at which h obtains its maximum, $f_{ii} = \pm 1$, $g_{ii} = \pm 1$, $h_{ii} = \pm 1$, and $e_{ii} = \pm 1$, and in addition f_{ii}, g_{ii}, h_{ii}, and e_{ii} satisfy the constraint $f_{ii}g_{ii}h_{ii}e_{ii} = 1$. For fixed i, the constraint is satisfied if either f_{ii}, g_{ii}, h_{ii}, and e_{ii} all have the same sign or exactly two of them are $+1$ and the other two are -1. Hence, for each fixed i there are $2 + \binom{4}{2} = 2^3$ choices of f_{ii}, g_{ii}, h_{ii}, and e_{ii} which satisfy the constraint. Since there are k values of i, there are 2^{3k} points at which h obtains its maximum.

Partition $O(k)$ into 2^k sets such that the signs of the diagonal elements of the matrices in any set are constant. This means that $T = (t_{ij}) \in O(k)$ and $S = (s_{ij}) \in O(k)$ are in the same set if and only if sign s_{ii} = sign t_{ii} $(1 \le i \le k)$. This procedure ignores matrices with diagonal elements equal to zero, but they constitute a set of measure 0 with respect to the invariant measure on $O(k)$. Now partition $V(k,p)$ and $V(k,q)$ into 2^k sets each, in a similar manner. Partitioning $O(k)$ (corresponding to F and G), $V(k,p)$ (corresponding to H_1), and $V(k,q)$ (corresponding to E_1), and forming the Cartesian product

$$\Lambda = V(k,p) \times O(k) \times D_V \times O(k) \times D_U \times V(k,q)$$

induces a partition of Λ into 2^{4k} disjoint sets. Exactly 2^{3k} of these sets contain in their interior one of the points at which h obtains its maximum. The other 2^k sets can be joined to these 2^{3k} sets to obtain a partition of Λ into 2^{3k} disjoint sets Λ_i $(1 \le i \le 2^{3k})$. It follows from the construction procedure that for each i, Ψh^n is integrable over Λ_i, h has an absolute maximum of

$$e^{-k} \prod_{i=1}^{k} (1-r_i\rho_i)^{-1}$$

on Λ_i, the maximum is obtained at an interior point, and Λ_i is of the form

$$\Lambda_i = M_{i1} \times M_{i2} \times D_V \times M_{i3} \times D_U \times M_{i4} \tag{3.2.6}$$

where $M_{i1} \subset V(k,p)$, $M_{i2} \subset O(k)$, $M_{i3} \subset O(k)$, and $M_{i4} \subset V(k,q)$.

Now write (3.2.5) as the sum of 2^{3k} integrals, where each integral is the integral of $C_n \Psi h^n$ over one of the Λ_i. The only part of the integrands which depend on F, G, H_1, or E_1 is λ, where $\lambda = \mathrm{etr}\{n[UFVGP_1H_1{}'R:O]E_1\}$. λ is invariant under transformations of the form

$$F \to FS_1,\ G \to S_1GS_2,\ H_1 \to \mathrm{diag}(S_3, I_{p-k})H_1S_2, \text{ and}$$

$$E_1 \to \mathrm{diag}(S_3, I_{q-k})E_1$$

where S_1, S_2, and S_3 are $k \times k$ diagonal matrices with ± 1 on the diagonal. The integral over Λ_i $(1 \le i \le 2^{3k})$ can be transformed by an appropriate choice of S_1, S_2, and S_3, where S_1, S_2, and S_3 depend on i, to an integral of the form

$$J_i = C_n \int_{\Omega_i} \Psi(y)\{h(y,\rho,r)\}^n dy\ , \tag{3.2.7}$$

where Ω_i is the imagine of Λ_i under the transformation. S_1, S_2, and S_3 are chosen so that the absolute maximum of h on Ω_i,

$$e^{-k} \prod_{i=1}^{k} (1-\rho_i r_i)^{-1} ,$$

is obtained at the interior point $\beta(\rho, r)$ defined by

$$(3.2.8) \qquad F = G = I_k , \quad H_1 = \begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix}, \quad E_1 = \begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix}, \quad \text{and}$$

$$u_i = v_i = (1-\rho_i r_i)^{-\frac{1}{2}} , \quad i=1,\ldots,k .$$

It follows from (3.2.6) that Ω_i is of the form

$$(3.2.9) \qquad \Omega_i = N_{i1} \times N_{i2} \times D_V \times N_{i3} \times D_U \times N_{i4}$$

where $N_{i1} \subset V(k,p)$, $N_{i2} \subset O(k)$, $N_{i3} \subset O(k)$ and $N_{i4} \subset V(k,q)$. For $m \geq k$ let $V^+(k,m)$ be the set of all $m \times k$ matrices in $V(k,m)$ with positive diagonal elements. If $k = m$ then write $O^+(k)$ for $V^+(k,m)$. It follows from the procedure used to construct the partition of Λ, that $V^+(k,p) \subset N_{i1}$, $O^+(k) \subset N_{i2}$, $O^+(k) \subset N_{i3}$, and $V^+(k,p) \in N_{i4}$. Combining these results we have

$$(3.2.10) \qquad {}_2F_1(n, P^2, R^2) = \sum_{i=1}^{2^{3k}} J_i .$$

We will now derive an asymptotic representation φ_i of J_i. In fact φ_i does not depend on i, i.e. $\varphi_i = \varphi$ for all i. It then follows from (3.2.10) that $2^{3k}\varphi$ is an asymptotic representation for ${}_2F_1(n, P^2, R^2)$.

Consider a particular J_i. Theorem 2.3.1 is not directly applicable to the integral representation (3.2.7) because the function h depends on $F, G \in O(k)$, $H_1 \in V(k,p)$, and $E_1 \in V(k,q)$ and the elements of these matrices are not independent since their columns are orthonormal.

It might be possible to generalize the theorem to handle this situation, but there would probably still be difficulties in computing a quantity which would correspond to the Hessian in the present statement of the theorem. The easiest approach seems to be to transform from F, G, H_1, and E_1 to a set of independent parameters. If $G \in O(k)$ then $G'G = I_k$ and hence $\det G = \pm 1$.

Definition 3.2.1--

$G \in O(k)$ is said to be proper if $\det G = 1$ and improper if $\det G = -1$.

Murnaghan (1938) proved that any proper $k \times k$ orthogonal matrix G can be expressed as $G = \exp(S)$ where

$$\exp(S) = I_k + S + \frac{S^2}{2!} + \frac{S^3}{3!} + \cdots$$

and $S = (s_{ij})$ is a $k \times k$ skew-symmetric matrix. The $\frac{1}{2}k(k-1)$ elements of S above the diagonal provide a parametrization of G. The mapping $G \to S$ defined by $G = \exp(S)$ is a mapping from $O(k)$ into $R^{\frac{1}{2}k(k-1)}$. The imagine of $O(k)$ under this mapping is a bounded subset of $R^{\frac{1}{2}k(k-1)}$ (c.f. Murnaghan (1938)).

Since $\det G$ is a continuous function of the elements of G and $\det I_k = 1$, there exists a neighborhood $N \subset O(k)$ of I_k such that $G \in N$ implies G is proper. It follows from (3.2.9) and the subsequent discussion, that we can choose $N \in O^+(k)$, $N_1 \in V^+(k,p)$, and $N_2 \in V^+(k,q)$ such that for all i $(1 \le i \le 2^{3k})$, $I_k \subset N \subset N_{i2}$, $N \subset N_{i3}$,

$$\begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix} \subset N_1 \subset N_{i1}, \quad \text{and} \quad \begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix} \subset N_2 \subset N_{i4}.$$

In particular, if $p = k$, take $N_1 = N$ and if $q = p = k$, take $N = N_1 = N_2$. Let

(3.2.11) $$\Xi = N_1 \times N \times D_V \times N \times D_U \times N_2 \ .$$

Then $\Xi \subset \Omega_i$, all orthogonal matrices in Ξ are proper, and $\beta(\rho, r)$ (defined by (3.2.8)) is an interior point of Ξ for all (ρ, r).

J_i may be written as

(3.2.12) $$J_i = C_n(L_1 + L_{i2})$$

where

(3.2.13) $$L_1 = \int_{\Xi} \eta(y)\{h(y, \rho, r)\}^n \, dy$$

and

(3.2.14) $$L_{i2} = \int_{\Omega_i - \Xi} \eta(y)\{h(y, \rho, r)\}^n \, dy \ .$$

L_1 and L_{i2} will be treated separately. We will use Theorem 2.3.1 to derive an asymptotic representation of L_1 and then we will prove that L_{i2} is asymptotically of lower order of magnitude than L_1.

Using the notation of Theorem 2.3.1, let

(3.2.15) $$A = \{(\rho, r) : 1 > \rho_1 > \rho_2 > \cdots > \rho_k > 0, \ 1 > r_1 > r_2 > \cdots > r_p > 0\}$$

and for each ϵ, $0 < \epsilon < \frac{1}{2}$, let $B = B(\epsilon)$ be the set of all $(\rho, r) \in A$ such that the ρ_i's differ from one another and 0 and 1 by at least ϵ and the r_i's differ from one another and 0 and 1 by at least ϵ. That is, if $\rho_0 = r_0 = 1$, $\rho_{k+1} = 0$, and $r_{p+1} = 0$, then

(3.2.16) $$B = \{(\rho, r) : \rho_i - \rho_{i+1} > \epsilon \text{ and } r_j - r_{j+1} > \epsilon \text{ for all } i(0 \le i \le k) \text{ and } j(0 \le j \le p)\} \ .$$

<u>Derivation of an Asymptotic Representation for L_1</u> --

F and G are proper in Ξ and hence can be parametrized by

(3.2.17) $$G = \exp(S), \quad S = (s_{ij}), \text{ and}$$

(3.2.18) $$F = \exp(T), \quad T = (t_{ij})$$

where S and T are $k \times k$ skew-symmetric matrices. Anderson (1965) computed the Jacobian of these transformations as

$$J(G \to S) = \frac{\Gamma_k(\frac{1}{2}k)}{2^k \pi^{\frac{1}{2}k^2}} \left\{ 1 + \frac{k-2}{24} \operatorname{tr} S^2 + \frac{8-k}{4\times 6!} \operatorname{tr} S^4 \right.$$

$$\left. + \frac{5k^2-20k+14}{8\times 6!} (\operatorname{tr} S^2)^2 + O(s_{ij}^{\ 5}) \right\}$$

with a similar expression for $J(F \to T)$. In this section the notation $O(s_{ij}^{\ m})$ means terms in the s_{ij} which are at least of order m. For our purposes it is sufficient to note that

$$J(G \to S) = \frac{\Gamma_k(\frac{1}{2}k)}{2^k \pi^{\frac{1}{2}k^2}} \{1 + O(s_{ij}^{\ 2})\}$$

and

$$J(F \to T) = \frac{\Gamma_k(\frac{1}{2}k)}{2^k \pi^{\frac{1}{2}k^2}} \{1 + O(t_{ij}^{\ 2})\} .$$

Let $[H_1 :-]$ be a $p \times p$ orthogonal matrix whose first k columns are H_1. James (1969) showed that a parametrization of H_1 is given by

$$[H_1 :-] = \exp(W) = \exp\left(\begin{bmatrix} W_{11} & W_{12} \\ -W_{12}' & 0 \end{bmatrix}\right) , \quad W = (w_{ij}) . \tag{3.2.19}$$

Similarly a parametrization of E_1 is given by

$$[E_1 :-] = \exp(Z) = \exp\left(\begin{bmatrix} Z_{11} & -Z_{21}' \\ Z_{21} & 0 \end{bmatrix}\right) , \quad Z = (z_{ij}) . \tag{3.2.20}$$

Here W_{11} and Z_{11} are $k \times k$ skew-symmetric matrices, W_{12} is $k \times (p-k)$, and Z_{21} is $(q-k) \times k$. The Jacobians of these transformations are (cf. James (1969))

$$J\{H_1 \to (W_{11}, W_{12})\} = \frac{\Gamma_k(\frac{1}{2}p)}{2^k \pi^{\frac{1}{2}kp}} \{1 + O(w_{ij}^{\ 2})\}$$

and

$$J\{E_1 \to (Z_{11}, Z_{21})\} = \frac{\Gamma_k(\tfrac{1}{2}q)}{2^k \pi^{\frac{1}{2}kq}} \{1 + O(z_{ij}^2)\} .$$

The point $\beta(\rho, r)$ defined by (3.2.8) is mapped into the point $\xi(\rho, r)$ defined by

(3.2.21) $$S = T = W_{11} = Z_{11} = O , \quad W_{12} = O, \quad Z_{21} = \Theta, \quad \text{and}$$
$$u_i = v_i = (1-\rho_i r_i)^{-\frac{1}{2}} \qquad (1 \le i \le k)$$

and Ξ is mapped into a set D containing $\xi(\rho, r)$ in its interior. It follows from (3.2.17)-(3.2.20) that

$$g_{ii} = 1 - \tfrac{1}{2} \sum_{j=1}^{k} s_{ij}^2 + O(s_{ij}^3) \qquad (1 \le i \le k)$$

$$g_{ij} = s_{ij} + O(s_{ij}^2) \qquad (1 \le i \ne j \le k) , \qquad s_{ij} = -s_{ji}$$

$$f_{ii} = 1 - \tfrac{1}{2} \sum_{j=1}^{k} t_{ij}^2 + O(t_{ij}^3) \qquad (1 \le i \le k)$$

$$f_{ij} = t_{ij} + O(t_{ij}^2) \qquad (1 \le i \ne j \le k) , \qquad t_{ij} = -t_{ji}$$

$$h_{ii} = 1 - \tfrac{1}{2} \sum_{j=1}^{p} w_{ij}^2 + O(w_{ij}^3) \qquad (1 \le i \le k)$$

$$h_{ij} = w_{ji} + O(w_{ij}^2) \qquad (1 \le i \le p, 1 \le j \le k, i \ne j) , \qquad w_{ij} = -w_{ji}$$

$$e_{ii} = 1 - \tfrac{1}{2} \sum_{j=1}^{q} z_{ij}^2 + O(z_{ij}^3) \qquad (1 \le i \le k) , \qquad \text{and}$$

$$e_{ij} = z_{ij} + O(z_{ij}^2) \qquad (1 \le i \le q, 1 \le j \le k) , \qquad z_{ij} = -z_{ji}$$

where $G = (g_{ij})$, $F = (f_{ij})$, $H_1 = (h_{ij})$, and $E_1 = (e_{ij})$. It follows from these relations after some simplification that

$$\mathrm{tr}\{[UFVGP_1H_1'R:O]E_1\} = \sum_{i=1}^{k} u_i v_i \rho_i r_i + \Psi$$

where

$$(3.2.22)\quad \Psi = -\sum_{i<j}^{k}\{\tfrac{1}{2}(u_iv_ir_i\rho_i+u_jv_jr_j\rho_j)(s_{ij}^{\,2}+t_{ij}^{\,2}+w_{ij}^{\,2}+z_{ij}^{\,2})$$

$$+(u_iv_j\rho_ir_i+u_jv_i\rho_jr_j)t_{ij}s_{ij}+(u_iv_j\rho_jr_i+u_jv_i\rho_ir_j)t_{ij}\,w_{ij}$$

$$+(u_iv_j\rho_jr_j+u_jv_i\rho_ir_i)t_{ij}z_{ij}+(u_iv_i\rho_jr_i+u_jv_j\rho_ir_j)s_{ij}\,w_{ij}$$

$$+(u_iv_i\rho_jr_j+u_jv_j\rho_ir_i)s_{ij}z_{ij}+(u_iv_i\rho_ir_j+u_jv_j\rho_jr_i)w_{ij}\,z_{ij}\}$$

$$-\sum_{i=1}^{k}\sum_{j=k+1}^{p}\{u_iv_i\rho_ir_jw_{ij}z_{ij}+\tfrac{1}{2}u_iv_i\rho_ir_i(w_{ij}^{\,2}+z_{ij}^{\,2})\}$$

$$-\tfrac{1}{2}\sum_{i=1}^{k}\sum_{j=p+1}^{q}u_iv_i\rho_ir_iz_{ij}^{\,2}\;.$$

Combining the previous results we have

$$(3.2.23)\qquad L_1 = \int_D \tau(x)\{f(x,\rho,r)\}^n dx$$

where

$$\tau(x) = \frac{\{\Gamma_k(\tfrac{1}{2}k)\}^2\Gamma_k(\tfrac{1}{2}p)\Gamma_k(\tfrac{1}{2}q)}{2^{4k}\pi^{k\{k+\frac{1}{2}(p+q)\}}}\prod_{i=1}^{k}(u_iv_i)^{-k}\prod_{i<j}^{k}\{(u_i^{\,2}-u_j^{\,2})(v_i^{\,2}-v_j^{\,2})\}$$

$$\times J(G\to S)J(F\to T)J\{H_1\to(W_{11},W_{12})\}J\{E_1\to(Z_{11},Z_{21})\}\;,$$

$$f(x,\rho,r) = \exp\{\sum_{i=1}^{k}(-\tfrac{1}{2}u_i^{\,2}-\tfrac{1}{2}v_i^{\,2}+u_iv_i\rho_ir_i)+\Psi\}\prod_{i=1}^{k}(u_iv_i)\;,$$

D is the image of Ξ, and $x \in D$.

An asymptotic representation of L_1 will now be obtained by applying Theorem 2.3.1 to (3.2.23). We must verify that the conditions of the theorem are satisfied. In the following discussion (i)-(ix) refer to the corresponding conditions (i)-(ix) of the theorem. Recall that A and $B = B(\epsilon)$ were defined by (3.2.15) and (3.2.16), respectively. $S(\delta,x)$, defined for all $\delta > 0$, is the open sphere with center x and radius δ.

Verification of the conditions of Theorem 2.3.1--

(i) $\tau(x)$ and $f(x,\rho,r)$ are non-negative functions on D. If $n > k-1$, then from (3.2.5)

$$ {}_2F_1(n,P^2,R^2) = C_n \int_\Lambda \eta(y)\{h(y,\rho,r)\}^n dy $$

where C_n and η are defined in (3.2.4) and h is defined in (3.2.5). $C_n > 0$ and η and h are non-negative on Λ and $\Lambda \times A$, respectively. Hence, for $n > k-1$

$$ C_n^{-1}\ {}_2F_1(n,P^2,R^2) \geq \int_\Xi \eta(y)\{h(y,\rho,r)\}^n dy $$

$$ = \int_D \tau(x)\{f(x,\rho,r)\}^n dx $$

where Ξ is defined in (3.2.11). Therefore $\tau(x)\{f(x,\rho,r)\}^n$ is absolutely integrable on D for all $(\rho,r) \in A$ and $n > k-1$. James (1968) shows that zonal polynomials on the domain of positive definite symmetric matrices are non-negative, increasing functions of each of the latent roots. The coefficients of the zonal polynomials in the series representations of ${}_2F_1(n,P^2,R^2)$ are positive because $n > k-1$ and $q \geq k$ (see 1.2.1 and 1.2.2). Therefore ${}_2F_1(n,P^2,R^2)$ is a non-negative increasing function of the ρ_i^2's and r_i^2's since P^2 and R^2 are positive definite. In particular

$$ {}_2F_1(n,P^2,R^2) \leq {}_2F_1\{n,(1-\epsilon)I_k,\ (1-\epsilon)I_k\} $$

for all $(\rho,r) \in B$. Take $r_n = C_n^{-1}\ {}_2F_1\{n,(1-\epsilon)I_k,\ (1-\epsilon)I_k\}$. Then

$$ r_n \geq \int_D |\tau(x)\{f(x,\rho,r)\}^n| dx $$

for all $(\rho,r) \in B$.

(ii) We have shown that for each $(\rho,r) \in A$, $f(x,\rho,r)$ has an absolute maximum value on D which is obtained at the interior point $\varsigma(\rho,r)$ defined by (3.2.21). From Lemma 3.2.2

$$\inf_B f\{\xi(\rho,r),\rho,r\} = \inf_B \{e^{-k} \prod_{i=1}^{k} (1-\rho_i r_i)^{-1}\}$$

$$> e^{-k} > 0 .$$

(iii) It follows from (3.2.11) that D is of the form

$$D = K_1 \times K_2 \times D_V \times K_3 \times D_U \times K_4 \tag{3.2.24}$$

where K_1 is a neighborhood of $(W_{11}=0,\ W_{12}=0)$, K_2 is a neighborhood of $(S=0)$, K_3 is a neighborhood of $(T=0)$, and K_4 is a neighborhood of $(Z_{11}=0,\ Z_{21}=0)$. Let $x_0(\rho,r)$ be the point $\{(1-\rho_1 r_1)^{-\frac{1}{2}},\ldots,(1-\rho_k r_k)^{-\frac{1}{2}}\}$. It follows from the definition of B that there exists a $\delta > 0$ such that $S\{\delta, x_0(\rho,r)\}$ is contained in the interior of D_V and D_U for every $(\rho,r) \in B$. This fact and (3.2.24) imply that there exists a δ_1 , $0 < \delta_1 \leq \delta$, such that $S\{\delta_1, \xi(\rho,r)\}$ is contained in the interior of D for every $(\rho,r) \in B$. $f(x,\rho,r) > 0$ for all $x \in S\{\delta_1,\xi(\rho,r)\}$ and $(\rho,r) \in B$ since $f(x,\rho,r) > 0$ for all $x \in D$ and $(\rho,r) \in A$.

(iv) For every $(\rho,r) \in A$, $h(y,\rho,r)$ is a continuously differentiable function of the components of $y \in \Lambda$, and (3.2.17)-(3.2.20) are continuously differentiable transformations from (F,G,H_1,E_1) to $(T,S,W_{11},W_{12},Z_{11},Z_{21})$. Therefore, for all $(\rho,r) \in A$,

$$\frac{\partial f}{\partial x_i}(x,\rho,r) \quad \text{and} \quad \frac{\partial^2 f}{\partial x_i \partial x_j}(x,\rho,r)$$

$(1 \leq i, j \leq k)$ exist and are continuous functions of x on D .

(v) We have to show that

$$0 < \inf_B \gamma_k^{\,2}(\rho,r) \quad \text{and} \quad \sup_B \gamma_1^{\,2}(\rho,r) < \infty$$

where $\gamma_1^{\,2}(\rho,r) \geq \gamma_2^{\,2}(\rho,r) \geq \cdots \geq \gamma_k^{\,2}(\rho,r)$ are the latent roots of $\Omega\{\xi(\rho,r),\rho,r\}$. $\Omega(x,\rho,r) = (\omega_{ij}(x,\rho,r))$ is the $k \times k$ matrix defined

for all $x \in D$ and $(\rho, r) \in A$ by

$$\omega_{ij}(x,\rho,r) = \frac{-\partial^2 \ln f}{\partial x_i \partial x_j}(x,\rho,r) .$$

To compute $\Omega\{\xi(\rho,r),\rho,r\}$ we need the second partial derivatives of $\psi(x,\rho,r) = -\ln f(x,\rho,r)$ evaluated at $\{\xi(\rho,r),\rho,r\}$. It follows from (3.2.22) that for fixed (ρ,r) , the only nonzero second partial derivatives of ψ evaluated at $\{\xi(\rho,r),\rho,r\}$ are

$$\frac{\partial^2\psi}{\partial u_i^2} = \frac{\partial^2\psi}{\partial v_i^2} = \rho_i r_i - 2 \qquad (1\le i\le k) ,$$

$$\frac{\partial^2\psi}{\partial u_i \partial v_i} = \rho_i r_i \qquad (1\le i\le k) ,$$

$$\frac{\partial^2\psi}{\partial s_{ij}^2} = \frac{\partial^2\psi}{\partial t_{ij}^2} = \frac{\partial^2\psi}{\partial w_{ij}^2} = \frac{\partial^2\psi}{\partial z_{ij}^2} = \frac{\rho_i r_i}{1-\rho_i r_i} + \frac{\rho_j r_j}{1-\rho_j r_j} \qquad (1\le i<j\le k) ,$$

$$\frac{\partial^2\psi}{\partial t_{ij}\partial s_{ij}} = \frac{\partial^2\psi}{\partial t_{ij}\partial z_{ij}} = \frac{\rho_i r_i + \rho_j r_j}{\{(1-\rho_i r_i)(1-\rho_j r_j)\}^{\frac{1}{2}}} \qquad (1\le i<j\le k) ,$$

$$\frac{\partial^2\psi}{\partial t_{ij}\partial w_{ij}} = \frac{\rho_j r_i + \rho_i r_j}{\{(1-\rho_i r_i)(1-\rho_j r_j)\}^{\frac{1}{2}}} \qquad (1\le i<j\le k) ,$$

$$\frac{\partial^2\psi}{\partial s_{ij}\partial z_{ij}} = \frac{\rho_j r_j}{1-\rho_i r_i} + \frac{\rho_i r_i}{1-\rho_j r_j} \qquad (1\le i<j\le k) ,$$

$$\frac{\partial^2\psi}{\partial w_{ij}\partial z_{ij}} = \frac{\rho_i r_j}{1-\rho_i r_i} + \frac{\rho_j r_i}{1-\rho_j r_j} \qquad (1\le i<j\le k) ,$$

$$\frac{\partial^2\psi}{\partial s_{ij}\partial w_{ij}} = \frac{\rho_j r_i}{1-\rho_i r_i} + \frac{\rho_i r_j}{1-\rho_j r_j} \qquad (1\le i<j\le k) ,$$

$$\frac{\partial^2\psi}{\partial w_{ij}^2} = \frac{\partial^2\psi}{\partial z_{ij}^2} = \frac{\rho_i r_i}{1-\rho_i r_i} \qquad (1\le i\le k,\ k+1\le j\le p) ,$$

$$\frac{\partial^2 \psi}{\partial w_{ij} \partial z_{ij}} = \frac{\rho_i r_j}{1-\rho_i r_i} \qquad (1 \le i \le k,\, k+1 \le j \le p) ,$$

and

$$\frac{\partial^2 \psi}{\partial z_{ij}^2} = \frac{\rho_i r_i}{1-\rho_i r_i} \qquad (1 \le i \le k,\, p+1 \le j \le q) .$$

It follows from these results and the fact that $(\rho, r) \in B$ implies $\epsilon < \rho_i,\ r_j < 1 - \epsilon \quad (1 \le i \le k,\ 1 \le j \le p)$ that

$$\sup_B |\omega_{ij}\{\xi(\rho,r),\rho,r\}| \le \max[2-\epsilon^2, 2(1-\epsilon)^2/\{\epsilon(2-\epsilon)\}] = \zeta < \infty .$$

By Lemma 2.3.1

$$\begin{aligned}
\gamma_1^2(\rho,r) &= \max_{x'x=1} x'\Omega\{\xi(\rho,r),\rho,r\}x \\
&= \max_{x'x=1} \sum_{i=1}^{k} \sum_{j=1}^{k} x_i x_j \omega_{ij}\{\xi(\rho,r),\rho,r\} \\
&\le \max_{x'x=1} \sum_{i=1}^{k} \sum_{j=1}^{k} |x_i||x_j||\omega_{ij}\{\xi(\rho,r),\rho,r\}| \\
&\le \zeta \max_{x'x=1} \left(\sum_{i=1}^{k} |x_i|\right)^2 \\
&< k^2 \zeta .
\end{aligned}$$

Let $\kappa^2 = k^2 \zeta$. Then

$$\sup_B \gamma_1^2(\rho,r) < \kappa^2 < \infty .$$

We will prove that

$$\inf_B \gamma_k^2(\rho,r) > 0$$

by computing $\Delta\{\xi(\rho,r),\rho,r\}$, the Hessian of ψ for fixed $(\rho,r) \in B$ evaluated at $\{\xi(\rho,r),\rho,r\}$. Since, by the same argument that was used

in the proof of Theorem 2.3.1, (ii)-(iv) imply that $\Omega\{\xi(\rho,r),\rho,r\}$ is positive definite for all $(\rho,r) \in A$, we have

$$\begin{aligned} (3.2.25) \qquad \Delta\{\xi(\rho,r)\} &= \det \Omega\{\xi(\rho,r),\rho,r\} \\ &= \prod_{i=1}^{k} \gamma_i^2(\rho,r) \\ &< \gamma_k^2(\rho,r)\varkappa^{2(k-1)} . \end{aligned}$$

We will prove that

$$0 < \chi = \inf_B \Delta\{\xi(\rho,r)\}$$

and this result together with (3.2.25) implies

$$\inf_B \gamma_k^2(\rho,r) \geq \varkappa^{-2(k-1)} \inf_B \Delta\{\xi(\rho,r)\} = \varkappa^{-2(k-1)} \chi > 0 .$$

The determinant of a matrix is unchanged if a pair of rows and the corresponding pair of columns are interchanged. By a sequence of such operations $\Delta\{\xi(\rho,r)\}$ can be reduced to the determinant of a block diagonal matrix of the form

$$\operatorname{diag}(W_1,\dots,W_k,X_{12},X_{13},\dots,X_{k-1,k},Y_{1,k+1},\dots,Y_{1p},\dots,Y_{k,k+1},\dots,Y_{kp},Z_1,\dots,Z_k)$$

where

$$W_i = \begin{bmatrix} \dfrac{\partial^2\psi}{\partial u_i^2} & \dfrac{\partial^2\psi}{\partial u_i \partial v_i} \\[2ex] \dfrac{\partial^2\psi}{\partial v_i \partial u_i} & \dfrac{\partial^2\psi}{\partial v_i^2} \end{bmatrix} \qquad (1 \leq i \leq k) ,$$

$$X_{ij} = \begin{bmatrix} \frac{\partial^2\psi}{\partial s_{ij}^2} & \frac{\partial^2\psi}{\partial s_{ij}\partial t_{ij}} & \frac{\partial^2\psi}{\partial s_{ij}\partial w_{ij}} & \frac{\partial^2\psi}{\partial s_{ij}\partial z_{ij}} \\ \frac{\partial^2\psi}{\partial t_{ij}\partial s_{ij}} & \frac{\partial^2\psi}{\partial t_{ij}^2} & \frac{\partial^2\psi}{\partial t_{ij}\partial w_{ij}} & \frac{\partial^2\psi}{\partial t_{ij}\partial z_{ij}} \\ \frac{\partial^2\psi}{\partial w_{ij}\partial s_{ij}} & \frac{\partial^2\psi}{\partial w_{ij}\partial t_{ij}} & \frac{\partial^2\psi}{\partial w_{ij}^2} & \frac{\partial^2\psi}{\partial w_{ij}\partial z_{ij}} \\ \frac{\partial^2\psi}{\partial z_{ij}\partial s_{ij}} & \frac{\partial^2\psi}{\partial z_{ij}\partial t_{ij}} & \frac{\partial^2\psi}{\partial z_{ij}\partial w_{ij}} & \frac{\partial^2\psi}{\partial z_{ij}^2} \end{bmatrix} \qquad (1\le i<j\le k),$$

$$Y_{ij} = \begin{bmatrix} \frac{\partial^2\psi}{\partial w_{ij}^2} & \frac{\partial^2\psi}{\partial w_{ij}\partial z_{ij}} \\ \frac{\partial^2\psi}{\partial z_{ij}\partial w_{ij}} & \frac{\partial^2\psi}{\partial z_{ij}^2} \end{bmatrix} \qquad (1\le i\le k, k+1\le j\le p), \quad \text{and}$$

$$Z_i = \frac{\rho_i r_i}{1-\rho_i r_i} I_{q-p} .$$

All derivatives are evaluated at the point $\{\xi(\rho,r),\rho,r\}$. We then have

$$\Delta\{\xi(\rho,r)\} = \prod_{i=1}^{k} \det W_i \prod_{i<j}^{k} \det X_{ij} \prod_{i=1}^{k} \prod_{j=k+1}^{p} \det Y_{ij} \prod_{i=1}^{k} \det Z_i .$$

After a considerable amount of simplification $\Delta\{\xi(\rho,r)\}$ reduces to

$$(3.2.26) \quad \Delta\{\xi(\rho,r)\} = 2^{2k} \prod_{i=1}^{k} \{(1-\rho_i r_i)^{p-q+1}(\rho_i r_i)^{q-p}\}$$

$$\times \prod_{i<j}^{k} \left\{ \frac{(\rho_i r_i-\rho_j r_j)^4(\rho_i^2-\rho_j^2)(r_i^2-r_j^2)}{(1-\rho_i r_i)^4(1-\rho_j r_j)^4} \right\} \prod_{i=1}^{k} \prod_{j=k+1}^{p} \left\{ \frac{\rho_i^2(r_i^2-r_j^2)}{(1-\rho_i r_i)^2} \right\} .$$

It follows from (3.2.26) and the fact that $(\rho,r) \in B$ implies $\epsilon < \rho_i, r_j < 1-\epsilon$ $(1\le i\le k,\ 1\le j\le p)$, that $\inf_B \Delta\{\xi(\rho,r)\} = \chi > 0$.

(vi) The set $D_1 = D_1(\epsilon)$ defined by $D_1(\epsilon) = \{\xi(\rho,r):(\rho,r) \in B\}$ is a bounded set contained in the interior of D. It follows from the definition of $\xi(\rho,r)$ (3.2.21) and the definition of D (3.2.24), that there exists a $\delta > 0$ such that $S(2\delta,x) \subset D$ for every $x \in D_1$. Let $D_2(\epsilon) = \{x:x \in S(\delta,\xi(\rho,r))$ for some $(\rho,r) \in B(\epsilon)\}$. Then $D_2(\epsilon)$ is a bounded set contained in the interior of D. Let $cl(D_2 \times B)$ be the closure of $D_2 \times B$. Since D_2 and B are bounded subsets of Euclidean space, $cl(D_2 \times B)$ is a closed, bounded, and hence compact, set. $\psi(x,\rho,r) = -\ln f(x,\rho,r)$ is a continuous function on $D \times A$ and hence, $\psi(x,\rho,r)$ is uniformly continuous on $cl(D_2 \times B) \subset D \times A$. In particular for any $\epsilon_1 > 0$ there exists a $\delta_3 > 0$ such that for all $(\rho,r) \in B$, $S\{\delta_3,\xi(\rho,r)\} \subset cl(D_2 \times B)$ and $x \in S\{\delta_3,\xi(\rho,r)\}$ implies $|\omega_{ij}(x,\rho,r) - \omega_{ij}\{\xi(\rho,r),\rho,r\}| < \epsilon_1$.

(vii) Recall that from (3.2.23) and (3.2.13)

$$L_1 = \int_D \tau(x)\{f(x,\rho,r)\}^n dx = \int_{\Xi} \eta(y)\{h(y,\rho,r)\}^n dy$$

where $y \to x$, $\Omega_i \supset \Xi \to D$, $h(y,\rho,r) \to f(x,\rho,r)$, $\eta(y) \to \tau(x)$, and $\beta(\rho,r) \to \xi(\rho,r)$ under the mapping

$$(F,G,H_1,E_1) \to (T,S,W_{11},W_{12},Z_{11},Z_{21}) .$$

Let q denote this mapping in terms of the x and y parameters, i.e., $x = q(y)$.

Condition (vii) of Theorem 2.3.1 says

a) for every $\delta_4 > 0$ there exists a constant θ, $0 < \theta < 1$, such that $|f(x,\rho,r)/f\{\xi(\rho,r),\rho,r\}| < \theta$ for all $x \in D - S\{\delta_4,\xi(\rho,r)\}$ and $(\rho,r) \in B$.

We will prove the following more general result

b) for every $\delta_4 > 0$ there exists a constant θ, $0 < \theta < 1$, such that $|h(y,\rho,r)/h\{\beta(\rho,r),\rho,r\}| < \theta$ for all $y \in \Omega_i - S\{\delta_4, \beta(\rho,r)\}$, $(\rho,r) \in B$, and $i(1 \leq i \leq 2^{3k})$.

The result b) will be used to prove that L_{i2}, defined by (3.2.14), is asymptotically of lower order of magnitude than L_1.

To prove that b) implies a), consider a fixed $\delta_4 > 0$. Since the set $D - S\{\delta_4, \xi(\rho,r)\}$ is closed relative to D and q is a continuous mapping from Ξ onto D, it follows that Π, the inverse image of $D - S\{\delta_4, \xi(\rho,r)\}$ under q, is closed relative to Ξ. Therefore $\Xi - \Pi$ is open relative to Ξ and, since $q\{\beta(\rho,r)\} = \xi(\rho,r)$, $\beta(\rho,r) \in \Xi - \Pi$. Hence, there exists a $\delta > 0$ such that $S\{\delta, \beta(\rho,r)\} \cap \Xi \subset \Xi - \Pi$. From b) there exists a constant θ, $0 < \theta < 1$, such that $|h(y,\rho,r)/h\{\beta(\rho,r),\rho,r\}| < \theta$ for all $y \in \Xi - S\{\delta, \beta(\rho,r)\}$ and $(\rho,r) \in B$. By definition $f(x,\rho,r) = h\{q^{-1}(x),\rho,r\}$ and $\Pi \subset \Xi - S\{\delta, \beta(\rho,r)\}$. Therefore $|f(x,\rho,r)/f\{\xi(\rho,r),\rho,r\}| < \theta$ for all $x \in D - S\{\delta_4, \xi(\rho,r)\}$ and $(\rho,r) \in B$.

Ω_i was defined in (3.2.9) as

$$\Omega_i = N_{i1} \times N_{i2} \times D_V \times N_{i3} \times D_U \times N_{i4}$$

where $N_{i1} \subset V(k,p)$, $N_{i2} \subset O(k)$, $N_{i3} \subset O(k)$, and $N_{i4} \subset V(k,q)$ are neighborhoods of

$$\begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix}, \quad I_k, \quad I_k, \quad \text{and} \quad \begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix},$$

respectively. We have used y to denote an arbitrary point in Ω_i, but in fact the components of y are values of u, v, G, F, H_1, and E_1. To emphasize this fact we will write

$$y = (H_1, G, v, F, u, E_1) .$$

For each point $y = (H_1, G, v, F, u, E_1) \in \Omega_i$, let

$$y_A = \left(\begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix}, \; I_k \,, \; v \,, \; I_k \,, \; u \,, \; \begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix} \right) .$$

Ω_i is a subset of Euclidean space so for $x, y \in \Omega_i$ let $d(x,y)$ denote the Euclidean distance between x and y, i.e., $d(x,y) = \{(x-y)'(x-y)\}^{\frac{1}{2}}$.

The following lemma is needed in the proof of b).

Lemma 3.2.3--

For every $\delta > 0$ there exists a constant $\theta = \theta(\delta)$, $0 < \theta < 1$, such that if $y \in \Omega_i$ for some i $(1 \leq i \leq 2^{3k})$ and $d(y_A, \beta(\rho, r)) \geq \delta$ for some $(\rho, r) \in B$, then

$$|h(y, \rho, r)/h\{\beta(\rho, r), \rho, r\}| < \theta .$$

Proof--

It follows from the definition of y_A and $\beta(\rho, r)$ that

$$d(y_A, \beta(\rho, r)) = \left[\sum_{i=1}^{k} [\{u_i - (1-\rho_i r_i)^{-\frac{1}{2}}\}^2 + \{v_i - (1-\rho_i r_i)^{-\frac{1}{2}}\}^2 \right]^{\frac{1}{2}} .$$

If $d(y_A, \beta(\rho, r)) \geq \delta$, then there exists a J $(1 \leq J \leq k)$ such that either

$$|u_J - (1-\rho_J r_J)^{-\frac{1}{2}}| \geq \delta/(2k)^{\frac{1}{2}} \tag{3.2.27}$$

or

$$|v_J - (1-\rho_J r_J)^{-\frac{1}{2}}| \geq \delta/(2k)^{\frac{1}{2}} . \tag{3.2.28}$$

From Lemma 3.2.2 and Theorem 2 of the appendix

$$\left|\frac{h(y, \rho, r)}{h\{\beta(\rho, r), \rho, r\}}\right| \leq \sup_{F, G, H_1, E_1} \left|\frac{h(y, \rho, r)}{h\{\beta(\rho, r), \rho, r\}}\right|$$

$$= \exp\left[\sum_{i=1}^{k} \{-\tfrac{1}{2}u_i^2 - \tfrac{1}{2}v_i^2 + \rho_i r_i u_i v_i + \ln(u_i v_i) + 1 + \ln(1-\rho_i r_i)\} \right]$$

$$\leq \exp\{g_1(u_J, v_J, \rho_J, r_J)\}$$

where $g_1(u_J, v_J, \rho_J, r_J) = -\frac{1}{2}u_J^2 - \frac{1}{2}v_J^2 + \rho_J r_J u_J v_J + \ln(u_J v_J) + 1 + \ln(1-\rho_J r_J)$. By Cauchy's inequality $v_J u_J \leq \frac{1}{2}(v_J^2 + u_J^2)$ and therefore $g_1(u_J, v_J, \rho_J, r_J) \leq g_2(u_J, v_J, \rho_J, r_J)$, where

$g_2(u_J, v_J, \rho_J, r_J) = -\frac{1}{2}u_J^2(1-\rho_J r_J) - \frac{1}{2}v_J^2(1-\rho_J r_J) + \ln(u_J v_J) + 1 + \ln(1-\rho_J r_J)$.

To prove the lemma it is sufficient to prove that

$$\sup_{v_J, u_J, \rho_J, r_J} g_2(u_J, v_J, \rho_J, r_J) = \nu < 0 .$$

The lemma will then follow by taking $\theta = e^{\nu}$. It is also sufficient to prove that

$$\sup_{v_J, u_J, \rho_J, r_J} g_1(u_J, v_J, \rho_J, r_J) < 0 .$$

However g_2 is much easier to work with than g_1 .

From (3.2.27) and (3.2.28) we have

$$u_J = (1-\rho_J r_J)^{-\frac{1}{2}} + d \cos w$$

$$v_J = (1-\rho_J r_J)^{-\frac{1}{2}} + d \sin w$$

where $d \geq \delta/(2k)^{\frac{1}{2}}$, $0 \leq w < 2\pi$, $d \cos w > -(1-\rho_J r_J)^{-\frac{1}{2}}$, and $d \sin w > -(1-\rho_J r_J)^{-\frac{1}{2}}$. Make this change of variables in g_2 . After simplifying we have

$$g_2(u_J, v_J, \rho_J, r_J) = g_3(d, w, \rho_J, r_J)$$

where

$$g_3(d, w, \rho_J, r_J) = -\frac{1}{2}d^2(1-\rho_J r_J) - d(1-\rho_J r_J)^{\frac{1}{2}}(\cos w + \sin w)$$

$$+ \ln\{1+d(1-\rho_J r_J)^{\frac{1}{2}}\cos w\} + \ln\{1+d(1-\rho_J r_J)^{\frac{1}{2}}\sin w\} .$$

A straightforward exercise in calculus leads to the result that for fixed d, ρ_J, and r_J

$$\sup_{w} g_3(d, w, \rho_J, r_J) = g_3(d, 0, \rho_J, r_J) = g_3(d, \tfrac{1}{2}\pi, \rho_J, r_J)$$

$$= g_4\{d(1-\rho_J r_J)^{\frac{1}{2}}\}$$

where

$$g_4(t) = -\tfrac{1}{2}t^2 - t + \ln(1+t) , \qquad t \geq 0 .$$

Since the location of the supremum of g_3 for fixed d, ρ_J, and r_J does not depend on d, ρ_J, and r_J, we have

$$\sup_{d, w, \rho_J, r_J} g_3(d, w, \rho_J, r_J) = \sup_{d(1-\rho_J r_J)^{\frac{1}{2}}} g_4\{d(1-\rho_J r_J)^{\frac{1}{2}}\} .$$

If $(\rho, r) \in B$ then $r_J, \rho_J < 1 - \epsilon$ and hence, since $\epsilon < \frac{1}{2}$, $(1-\rho_J r_J)^{\frac{1}{2}} \geq \{1 - (1-\epsilon)^2\}^{\frac{1}{2}} = (2\epsilon - \epsilon^2)^{\frac{1}{2}} \geq \epsilon^{\frac{1}{2}}$. Also $g_4(t)$ is a decreasing function for $t \geq 0$ and $g(0) = 0$. Combining these results we have

$$\sup_{d(1-\rho_J r_J)^{\frac{1}{2}}} g_4\{d(1-\rho_J r_J)^{\frac{1}{2}}\} \leq g_4\{\epsilon^{\frac{1}{2}}\delta/(2k)^{\frac{1}{2}}\} < g(0) = 0$$

proving the lemma.

If $(\rho, r) \in B$ then for all j $(1 \leq j \leq k-1)$

$$\inf_{B} \left| (1-\rho_j r_j)^{-\frac{1}{2}} - (1-\rho_{j+1} r_{j+1})^{-\frac{1}{2}} \right| = \left| \{1-(2\epsilon)^2\}^{-\frac{1}{2}} - (1-\epsilon^2)^{-\frac{1}{2}} \right| = \chi(\epsilon) > 0 .$$

Let $\delta_4 > 0$ be given. Without loss of generality we may assume that $\delta_4 \leq \min\{\epsilon, \frac{1}{2}\chi(\epsilon)\}$. Choose i $(1 \leq i \leq 2^{3k})$ and $(\rho, r) \in B$. If $y \in \Omega_i - S\{\delta_4, \beta(\rho, r)\}$ then $d(y, \beta(\rho, r)) \geq \delta_4$. By the triangle inequality

$$d(y, \beta(\rho, r)) \leq d(y, y_A) + d(y_A, \beta(\rho, r))$$

and thus $d(y, \beta(\rho, r) \geq \delta_4$ implies $d(y_A, \beta(\rho, r)) \geq \frac{1}{2}\delta_4$ or $d(y_A, y) \geq \frac{1}{2}\delta_4$. If $d(y_A, \beta(\rho, r)) \geq \frac{1}{2}\delta_4$ then it follows from Lemma 3.2.3 with $\delta = \frac{1}{2}\delta_4$, that there exists a $\theta_1 = \theta_1(\delta_4)$, $0 < \theta_1 < 1$, such that $|h(y, \rho, r)/h\{\beta(\rho, r), \rho, r\}| < \theta_1$. Therefore we need only consider those y such that $d(y_A, \beta(\rho, r)) < \frac{1}{2}\delta_4$ and $d(y_A, y) \geq \frac{1}{2}\delta_4$.

The sets $O(k)$, $\mathcal{V}(k,p)$, and $\mathcal{V}(k,q)$ are bounded subsets of $R^{\frac{1}{2}k(k-1)}$, $R^{\frac{1}{2}k(2p-k-1)}$, and $R^{\frac{1}{2}k(2q-k-1)}$, respectively. Consequently N_{i1}, N_{i2}, N_{i3}, and N_{i4} , which occur in the definition of Ω_i , are bounded. However Ω_i is not bounded because D_V and D_U are not bounded. We want to show that the set of all $y \in \Omega_i$ satisfying $d(y_A, \beta(\rho,r)) < \frac{1}{2}\delta_4$ for some $(\rho,r) \in B$ is contained in a closed bounded subset Δ_i of Ω_i , and furthermore, that the u_i's and v_i's of the y's in Δ_i are bounded away from one another and from 0 and 1. The properties of Δ_i will be developed in the following discussion.

If $(\rho,r) \in B$ then for all j $(1 \leq j \leq k)$

$$1 < (1-\rho_j r_j)^{-\frac{1}{2}} \leq \epsilon^{-\frac{1}{2}} \leq \delta_4^{-\frac{1}{2}} .$$

At the point $\beta(\rho,r)$, $u_j = v_j = (1-\rho_j r_j)^{-\frac{1}{2}}$ $(1 \leq j \leq k)$. Hence $d(y_A, \beta(\rho,r)) < \frac{1}{2}\delta_4$ implies

$$\left[\sum_{j=1}^{k} [\{u_j-(1-\rho_j r_j)^{-\frac{1}{2}}\}^2 + \{v_j-(1-\rho_j r_j)^{-\frac{1}{2}}\}^2]\right]^{\frac{1}{2}} < \tfrac{1}{2}\delta_4$$

and thus $|u_j - (1-\rho_j r_j)^{-\frac{1}{2}}| < \frac{1}{2}\delta_4$ and $|v_j - (1-\rho_j r_j)^{-\frac{1}{2}}| < \frac{1}{2}\delta_4$ for all j . Combining these results we have after some simplification

$$\delta_4 \leq \epsilon < u_j,\ v_j < \epsilon + \epsilon^{-\frac{1}{2}} < 2\delta_4^{-\frac{1}{2}}$$

for all j . Also

$$\begin{aligned}
2\delta_4 \leq \chi(\epsilon) &\leq |(1-\rho_j r_j)^{-\frac{1}{2}}-(1-\rho_{j+1} r_{j+1})^{-\frac{1}{2}}| \\
&\leq |(1-\rho_j r_j)^{-\frac{1}{2}}-u_j| + |u_j-u_{j+1}| + |u_{j+1}-(1-\rho_{j+1} r_{j+1})^{-\frac{1}{2}}| \\
&< \tfrac{1}{2}\delta_4 + |u_j-u_{j+1}| + \tfrac{1}{2}\delta_4 .
\end{aligned}$$

Therefore $|u_j-u_{j+1}| > \delta_4$ $(1 \leq j \leq k-1)$. Similarly $|v_j-v_{j+1}| > \delta_4$ $(1 \leq j \leq k-1)$. Let $D_Z(\delta_4)$ be the set defined by

$$D_Z(\delta_4) = \{(z_1,\ldots,z_k): 2\delta_4^{-\frac{1}{2}} > z_1 > z_2 > \cdots > z_k > \delta_4 \text{ and }$$
$$z_j - z_{j+1} > \delta_4 \quad (1 \le j \le k-1)\} .$$

It follows from the previous discussion that if $y \in \Omega_i$ and $d(y_A, \beta(\rho,r)) < \frac{1}{2}\delta_4$ for some $(\rho,r) \in B$, then $y \in \Delta_i$ where

$$\Delta_i = N_{i1} \times N_{i2} \times D_V(\delta_4) \times N_{i3} \times D_U(\delta_4) \times N_{i4} .$$

Δ_i is bounded and $\beta(\rho,r)$ is an interior point of Δ_i for every $(\rho,r) \in B$. Let

$$\Gamma_i = \{y \in \Omega_i : d(y, y_A) \ge \tfrac{1}{2}\delta_4\}$$

and let $\Psi_i = \Delta_i \cap \Gamma_i$. Then Ψ_i is bounded. $B = B(\epsilon)$ is also bounded and therefore $\mathrm{cl}(\Psi_i \times B)$, the closure of $\Psi_i \times B$, is closed and bounded, and hence compact. Let $h_1(y,\rho,r)$ be the function defined on $\mathrm{cl}(\Psi_i \times B)$ by

$$h_1(y,\rho,r) = \frac{h(y,\rho,r)}{h\{\beta(\rho,r),\rho,r\}} .$$

h_1 is non-negative and continuous. Since $(y,\rho,r) \in \mathrm{cl}(\Psi_i \times B)$ implies $u_1 > u_2 > \cdots > u_k > 0$ and $v_1 > v_2 > \cdots > v_k > 0$, it follows from Lemma 3.2.2 that $h_1 < 1$. Hence from a well known result (see for example Goffman (1965), p. 21), there exists a $\theta(i) < 1$ such that $h_1(y,\rho,r) < \theta(i)$ for all $(y,\rho,r) \in \mathrm{cl}(\Psi_i \times B)$. Let $\theta = \max\{\theta_1, \theta(1), \theta(2), \ldots, \theta(k)\}$. Then we have proved

<u>Lemma 3.2.4</u>--

For every $\delta_4 > 0$ there exists a constant $\theta = \theta(\delta_4)$, $0 < \theta < 1$, such that

$$\left| \frac{h(y,\rho,r)}{h\{\beta(\rho,r),\rho,r\}} \right| < \theta$$

for all $y \in \Omega_i - S\{\delta_4, \beta(\rho,r)\}$, $(\rho,r) \in B$, and i $(1 \le i \le 2^{3k})$.

(viii) From (3.2.23) we have

$$(3.2.29) \qquad \tau\{\xi(\rho,r)\} = C_1 \prod_{i=1}^{k} (1-\rho_i r_i)^k \prod_{i<j}^{k} \left\{ \frac{\rho_i r_i - \rho_j r_j}{(1-\rho_i r_i)(1-\rho_j r_j)} \right\}^2$$

where

$$C_1 = \frac{\{\Gamma_k(\tfrac{1}{2}k)\}^2 \Gamma_k(\tfrac{1}{2}p)\Gamma_k(\tfrac{1}{2}q)}{2^{4k}\pi^{k\{k+\frac{1}{2}(p+q)\}}} .$$

Since $(\rho,r) \in B$ implies $\epsilon < \rho_i, \ r_i < 1-\epsilon \quad (1 \leq i \leq k)$,

$$(3.2.30) \qquad \inf_B \tau\{\xi(\rho,r)\} \geq C_1 \prod_{i=1}^{k} \{1-(1-\epsilon)^2\}^k \prod_{i<j}^{k} \left\{ \frac{3\epsilon^2}{(1-\epsilon^2)^2} \right\}^2 > 0 .$$

(ix) Let $D_2(\epsilon) = \{x : x \in S(\delta, \xi(\rho,r))\}$ be the subset of D defined in (vi). $D_2(\epsilon)$ is bounded and therefore $\text{cl}\{D_2(\epsilon)\} \subset D$ is compact. τ is continuous on D and hence uniformly continuous on $\text{cl}\{D_2(\epsilon)\}$. In particular, for any $\epsilon_1 > 0$ there exists a $\delta_5 > 0$ such that for all $(\rho,r) \in B$, $S\{\delta_5, \xi(\rho,r)\} \subset \text{cl}\{D_2(\epsilon)\}$ and $x \in S\{\delta_5, \xi(\rho,r)\}$ implies $|\tau(x) - \tau\{\xi(\rho,r)\}| < \epsilon_1$.

All of the conditions of Theorem 2.3.1 are satisfied and therefore

$$(3.2.31) \qquad L_1 \sim \tau\{\xi(\rho,r)\}[f\{\xi(\rho,r),\rho,r\}]^n[\Delta\{\xi(\rho,r)\}]^{-\frac{1}{2}}(2\pi/n)^{\frac{1}{2}k(p+q)}$$

uniformly in $(\rho,r) \in B$. Substituting (3.2.26) and (3.2.29) into (3.2.31) and simplifying gives

$$(3.2.32) \qquad L_1 \sim C_2 \prod_{i=1}^{k} \{(1-\rho_i r_i)^{-n+\frac{1}{2}(q+p-1)}(\rho_i r_i)^{\frac{1}{2}(p-q)}\}$$

$$\times \prod_{i<j}^{k} \{(\rho_i^2-\rho_j^2)(r_i^2-r_j^2)\}^{-\frac{1}{2}} \prod_{i=1}^{k} \prod_{j=k+1}^{p} \{\rho_i^2(r_i^2-r_j^2)\}^{-\frac{1}{2}}$$

uniformly in $(p,r) \in B$, where

$$C_2 = \frac{\{\Gamma_k(\tfrac{1}{2}k)\}^2 \Gamma_k(\tfrac{1}{2}p)\Gamma_k(\tfrac{1}{2}q)e^{-nk}}{n^{\frac{1}{2}k(p+q)}\, 2^{5k-\frac{1}{2}k(p+q)}\, \pi^{k^2}} .$$

Asymptotic behavior of L_{i2}--

It follows from (3.2.14) that

$$L_{i2} = \int_{\Omega_i - \Xi} \eta(y)\{h(y,\rho,r)\}^n dy .$$

We can show, by exactly the same argument that was used in (i) to prove the absolute integrability of $\tau(x)\ \{f(x,\rho,r)\}^n$ on D, that $\eta(y)\ \{h(y,\rho,r)\}^n$ is integrable on $\Omega_i - \Xi$ for all i $(1 \le i \le 2^{3k})$ and $n > k-1$. Furthermore

$$r_n \ge \int_{\Omega_i - \Xi} \eta(y)\ \{h(y,\rho,r)\}^n dy$$

for all $n > k-1$, $(\rho,r) \in B$, and i $(1 \le i \le 2^{3k})$, where $r_n = C_n^{-1}\ {}_2F_1(n,(1-\epsilon)I_k,\ (1-\epsilon)I_k)$ and C_n is defined by (3.2.4). Integrability of ηh^n is the same as absolute integrability since η and h are positive on Ω_i and $\Omega_i \times A$, respectively. It follows from the definition of Ξ (3.2.11) and Lemma 3.2.4, that there exists a constant θ, $0 < \theta < 1$, such that $|h(y,\rho,r)/h\{\beta(\rho,r),\rho,r\}| < \theta$ for all $y \in \Omega_i - \Xi$, $(\rho,r) \in B$, and i $(1 \le i \le 2^{3k})$. Therefore

$$\begin{aligned}
|L_{i2}| &= \int_{\Omega_i - \Xi} |\eta(y)\{h(y,\rho,r)\}^n|\, dy \\
&\le \int_{\Omega_i - \Xi} |\eta(y)\{h(y,\rho,r)\}^k|\ \left|\frac{h(y,\rho,r)}{h\{\beta(\rho,r),\rho,r\}}\right|^{n-k} |h\{\beta(\rho,r),\rho,r\}|^{n-k} dy \\
&\le \theta^{n-k} |h\{\beta(\rho,r),\rho,r\}|^{n-k} \int_{\Omega_i - \Xi} |\eta(y)\{h(y,\rho,r)\}^k|\, dy \\
&\le \theta^{n-k} r_k |h\{\beta(\rho,r),\rho,r\}|^{n-k} .
\end{aligned}$$

We have

$$h\{\beta(\rho,r),\rho,r\} = f\{\xi(\rho,r),\rho,r\} = e^{-k} \prod_{i=1}^{k} (1-\rho_i r_i)^{-1} . \tag{3.2.33}$$

It follows from (3.2.33), (3.2.26), and (3.2.30) that

$$\sup_B [\tau\{\xi(\rho,r)\}^{-1}[f\{\xi(\rho,r),\rho,r\}]^{-k}[\Delta\{\xi(\rho,r)\}]^{\frac{1}{2}}(2\pi)^{-\frac{1}{2}k(p+q)}] = \iota < \infty .$$

Let $\hat{L}_1$ denote the righthand side of (3.2.31). Then

$$|L_{i2}/\hat{L}_1| \leq \theta^{n-k} n^{\frac{1}{2}k(p+q)} r_k \iota \to 0 \qquad \text{as } n \to \infty$$

uniformly in $(\rho,r) \in B$ and i $(1 \leq i \leq 2^{3k})$. Hence L_{i2} is asymptotically of lower order of magnitude than L_1. Consequently

$$L_1 + L_{i2} \sim \hat{L}_1 . \tag{3.2.34}$$

Combining (3.2.10), (3.2.12), and (3.2.34), we have

Theorem 3.2.1--

Let

i) $R = \mathrm{diag}(r_1,\ldots,r_p)$ with $1 > r_1 > \cdots > r_p > 0$;

ii) $P = \mathrm{diag}(\rho_1,\ldots,\rho_p)$ with $1 > \rho_1 > \cdots > \rho_k > \rho_{k+1} = \cdots = \rho_p = 0$;

iii) $A = \{(\rho,r): 1 > \rho_1 > \cdots > \rho_k > 0 ,\ 1 > r_1 > \cdots > r_p > 0\}$ where $\rho = (\rho_1,\ldots,\rho_k)'$ and $r = (r_1,\ldots,r_p)'$; and

iv) for every ϵ, $0 < \epsilon < \frac{1}{2}$,

$$B = B(\epsilon) = \{(\rho,r) \in A : \rho_i - \rho_{i+1} > \epsilon \ (0 \leq i \leq k) ,\ r_j - r_{j+1} > \epsilon \ (0 \leq j \leq p) ,\ \text{where } \rho_0 = r_0 = 1 ,\ \rho_{k+1} = r_{p+1} = 0\} .$$

Then for all $(\rho,r) \in A$

$$ {}_2F_1(\tfrac{1}{2}n,\tfrac{1}{2}n;\tfrac{1}{2}q;P^2,R^2) \sim K_n \prod_{i=1}^{k} \{(1-\rho_i r_i)^{-n+\frac{1}{2}(q+p-1)}(\rho_i r_i)^{\frac{1}{2}(p-q)}\} \times \prod_{i=1}^{k}\prod_{j=1}^{p} \{(r_i^2-r_j^2)(\rho_i^2-\rho_j^2)\}^{-\frac{1}{2}} \tag{3.2.35}$$

where

$$K_n = \frac{(\frac{1}{2}n)^{nk-\frac{1}{2}k(p+q)} e^{-nk} \Gamma_k(\frac{1}{2}p)\Gamma_k(\frac{1}{2}q)}{\{\Gamma_k(\frac{1}{2}n)\}^2} .$$

Furthermore (3.2.35) holds uniformly in $(\rho, r) \in B$ for every ϵ .

As a check on the preceding derivation it should be noted that when $p = k = 1$, (3.2.35) agrees with the known asymptotic behavior of the classical hypergeometric function (see Luke (1969, Section 7.21)).

3.3. Partial Differential Equations for Hypergeometric Functions of Two Matrix Arguments.

In this section the partial differential equations derived by Constantine and Muirhead (1972) are extended to include the case where some of the latent roots of one of the matrices are known to be zero. These results and the results of Section 3.2 will be used in the next section to find additional terms in the asymptotic expansion of ${}_2F_1(\mathfrak{n}, P^2, R^2)$.

We introduce the following notation:

1) S is a $p \times p$ symmetric matrix with latent roots $s_1, \ldots, s_p$;
2) $T = \mathrm{diag}(T_1, 0)$ is a $p \times p$ symmetric matrix and T_1 is an $\ell \times \ell$ symmetric matrix with latent roots $t_1, \ldots, t_\ell$;
3) $\kappa = (k_1, \ldots, k_p)$, $k_1 \geq k_2 \geq \cdots \geq k_p \geq 0$, is a partition of $k = \Sigma_{i=1}^p k_i$ into p parts;
4) $\eta = (k_1, \ldots, k_\ell)$, $k_1 \geq \cdots \geq k_\ell \geq 0$ is a partition defined by the partition κ , of $h = \Sigma_{i=1}^{\ell} k_i$ into ℓ parts;
5) Corresponding to the partition κ , let
   $\kappa_i = (k_1, \ldots, k_{i-1}, k_i + 1, k_{i+1}, \ldots, k_p)$ and
   $\kappa^{(i)} = (k_1, \ldots, k_{i-1}, k_i - 1, k_{i+1}, \ldots, k_p)$ whenever they are admissible, i.e., whenever the elements of the partition are nonnegative and in nonincreasing order;
6) Corresponding to the partition η , let
   $\eta_i = (k_1, \ldots, k_{i-1}, k_i + 1, k_{i+1}, \ldots, k_\ell)$ and

$\eta^{(i)} = (k_1, \ldots, k_{i-1}, k_i - 1, k_{i+1}, \ldots, k_\ell)$ whenever they are admissible; and

7) If $\tau = (k_1, \ldots, k_\ell)$ then $C_\tau(B)$ will denote the zonal polynomial of the $p \times p$ matrix B corresponding to the partition $(k_1, \ldots, k_\ell, 0, \ldots, 0)$.

James (1964) proved that

$$C_\varkappa(T) = \begin{cases} C_\eta(T_1) & \text{if} \quad k_{\ell+1} = \cdots = k_p = 0 \\ 0 & \text{otherwise.} \end{cases} \tag{3.3.1}$$

From (3.3.1) we have

$$\sum_{k=0}^{\infty} \sum_{\varkappa} \frac{(a_1)_\varkappa \cdots (a_u)_\varkappa}{(b_1)_\varkappa \cdots (b_v)_\varkappa} \frac{C_\varkappa(T) C_\varkappa(S)}{k!\, C_\varkappa(I_p)} \tag{3.3.2}$$

$$= \sum_{h=0}^{\infty} \sum_{\eta} \frac{(a_1)_\eta \cdots (a_u)_\eta}{(b_1)_\eta \cdots (b_v)_\eta} \frac{C_\eta(T_1) C_\eta(S)}{h!\, C_\eta(I_\ell)}$$

where $a_1, \ldots, a_u$, $b_1, \ldots, b_v$ are real or complex constants. We need the following differential operators

$$\delta_S = \sum_{i=1}^{p} s_i \frac{\partial^2}{\partial s_i^2} + \sum_{i=1}^{p} \sum_{\substack{j=1 \\ i \neq j}}^{p} \frac{s_i}{s_i - s_j} \frac{\partial}{\partial s_i} , \tag{3.3.3}$$

$$\epsilon_S = \sum_{i=1}^{p} \frac{\partial}{\partial s_i} , \tag{3.3.4}$$

$$\gamma_T = \sum_{i=1}^{\ell} t_i^2 \frac{\partial}{\partial t_i} , \tag{3.3.5}$$

and

$$\eta_T = \sum_{i=1}^{\ell} t_i^3 \frac{\partial^2}{\partial t_i^2} + \sum_{i=1}^{\ell} \sum_{\substack{j=1 \\ i \neq j}}^{\ell} \frac{t_i^3}{t_i - t_j} \frac{\partial}{\partial t_i} + (1 - \tfrac{1}{2}(\ell - 1))\gamma_T . \tag{3.3.6}$$

Constantine and Muirhead (1972) prove

(3.3.7) $$\delta_S C_\kappa(S) = C_\kappa(I_p) \sum_{i=1}^{p} \binom{\kappa}{\kappa(i)} [k_i - 1 + \tfrac{1}{2}(p-i)] \frac{C_{\kappa(i)}(S)}{C_{\kappa(i)}(I_p)}$$

(3.3.8) $$\epsilon_S C_\kappa(S) = C_\kappa(I_p) \sum_{i=1}^{p} \binom{\kappa}{\kappa(i)} \frac{C_{\kappa(i)}(S)}{C_{\kappa(i)}(I_p)}$$

(3.3.9) $$\gamma_T C_\eta(T_1) = \frac{1}{h+1} \sum_{i=1}^{\ell} \binom{\eta_i}{\eta} [k_i - \tfrac{1}{2}(i-1)] C_{\eta_i}(T_1)$$

(3.3.10) $$\eta_T C_\eta(T_1) = \frac{1}{h+1} \sum_{i=1}^{\ell} \binom{\eta_i}{\eta} [k_i - \tfrac{1}{2}(i-1)][k_i - \tfrac{1}{2}(i-\ell)] C_{\eta_i}(T_1)$$

and

(3.3.11) $$\sum_{i=1}^{\ell} \binom{\eta_i}{\eta} C_{\eta_i}(T_1) = (h+1)\mathrm{Tr}(T_1) C_\eta(T_1)$$

where $\binom{\kappa}{\sigma}$ is the coefficient of $C_\sigma(S)/C_\sigma(I_p)$ in the "binomial expansion"

$$\frac{C_\kappa(I_p+S)}{C_\kappa(I_p)} = \sum_{j=0}^{k} \sum_{\sigma} \binom{\kappa}{\sigma} \frac{C_\sigma(S)}{C_\sigma(I_p)}$$

and $\binom{\eta_i}{\tau}$ is the coefficient of $C_\tau(T_1)/C_\tau(I_\ell)$ in the "binomial expansion"

$$\frac{C_{\eta_i}(I_\ell+T_1)}{C_{\eta_i}(I_\ell)} = \sum_{j=0}^{\ell} \sum_{\tau} \binom{\eta_i}{\tau} \frac{C_\tau(T_1)}{C_\tau(I_\ell)}$$

(see Muirhead (1970) for details).

We are now in a position to extend the partial differential equations satisfied by hypergeometric functions commonly occurring in multivariate analysis. We start with the ${}_2F_1$ function

Theorem 3.3.1--

The function ${}_2F_1(a,b;c;T,S)$ satisfies the partial differential equation

(3.3.12) $$\delta_S F + [c - \tfrac{1}{2}(p-1)]\epsilon_S F - [a+b-\tfrac{1}{2}(\ell-1)]\gamma_T F - \eta_T F = ab\,\mathrm{Tr}(T_1)F\ .$$

Proof--

$$(3.3.13) \qquad {}_2F_1(a,b;c;T,S) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a)_\kappa (b)_\kappa}{(c)_\kappa} \frac{C_\kappa(S) C_\kappa(T)}{k!\, C_\kappa(I_p)}$$

By applying the differential operator δ_S term-by-term to the series (3.3.13) it follows from (3.3.7) and (3.3.1) that

$$(3.3.14) \quad \delta_S\, {}_2F_1 = \sum_{h=0}^{\infty} \sum_{\eta} \frac{(a)_\eta (b)_\eta}{(c)_\eta\, h!} C_\eta(T_1) \sum_{i=1}^{\ell} \binom{\eta}{\eta(i)} [k_i - 1 + \tfrac{1}{2}(p-i)] \frac{C_{\eta(i)}(S)}{C_{\eta(i)}(I_p)} .$$

By applying ϵ_S it follows from (3.3.8) and (3.3.1) that

$$(3.3.15) \qquad \epsilon_S\, {}_2F_1 = \sum_{h=0}^{\infty} \sum_{\eta} \frac{(a)_\eta (b)_\eta}{(c)_\eta\, h!} C_\eta(T_1) \sum_{i=1}^{\ell} \binom{\eta}{\eta(i)} \frac{C_{\eta(i)}(S)}{C_{\eta(i)}(I_p)} .$$

Using (3.3.2) we may write the ${}_2F_1$ function as

$$(3.3.16) \qquad {}_2F_1(a,b;c;T,S) = \sum_{h=0}^{\infty} \sum_{\eta} \frac{(a)_\eta (b)_\eta}{(c)_\eta\, h!} \frac{C_\eta(T_1) C_\eta(S)}{C_\eta(I_p)} .$$

By applying the operator γ_T to the series (3.3.16) it follows from (3.3.9) that

$$(3.3.17) \quad \gamma_T\, {}_2F_1 = \sum_{h=0}^{\infty} \sum_{\eta} \frac{(a)_\eta (b)_\eta}{(c)_\eta (h+1)!} \frac{C_\eta(S)}{C_\eta(I_p)} \sum_{i=1}^{\ell} \binom{\eta_i}{\eta} [k_i - \tfrac{1}{2}(i-1)] C_{\eta_i}(T_1) .$$

By applying η_T to (3.3.16) it follows from (3.3.10) that

$$(3.3.18) \quad \eta_T\, {}_2F_1 =$$

$$\sum_{h=0}^{\infty} \sum_{\eta} \frac{(a)_\eta (b)_\eta}{(c)_\eta (h+1)!} \frac{C_\eta(S)}{C_\eta(I_p)} \sum_{i=1}^{\ell} \binom{\eta_i}{\eta} [k_i - \tfrac{1}{2}(i-1)][k_i - \tfrac{1}{2}(i-\ell)] C_{\eta_i}(T_1) .$$

Let Δ denote the differential operator

$$\delta_S + [c - \tfrac{1}{2}(p-1)]\epsilon_S - [a+b-\tfrac{1}{2}(\ell-1)]\gamma_T - \eta_T$$

Combining (3.3.14)-(3.3.18) we have

$$(3.3.19)\quad \Delta_2 F_1 = \sum_{h=0}^{\infty} \sum_{\eta} \frac{(a)_\eta (b)_\eta}{(c)_\eta\, h!} \left\{ C_\eta(T_1) \sum_{i=1}^{\ell} \binom{\eta}{\eta^{(i)}} \frac{C_{\eta^{(i)}}(S)}{C_{\eta^{(i)}}(I_p)} [k_i + c - \tfrac{1}{2}(i+1)] \right.$$

$$\left. - \frac{C_\eta(S)}{(h+1) C_\eta(I_p)} \sum_{i=1}^{\ell} \binom{\eta_i}{\eta} C_{\eta_i}(T_1) [k_i - \tfrac{1}{2}(i-1)][a+b+k_i - \tfrac{1}{2}(i-1)] \right\} .$$

Let C denote the coefficient of $C_\eta(S)$ in (3.3.19). Then

$$C = \sum_{i=1}^{\ell} \binom{\eta_i}{\eta} \frac{C_{\eta_i}(T_1)}{(h+1)!\, C_\eta(I_p)} \left\{ \frac{(a)_{\eta_i} (b)_{\eta_i}}{(c)_{\eta_i}} [k_i + c - \tfrac{1}{2}(i-1)] \right.$$

$$\left. - \frac{(a)_\eta (b)_\eta}{(c)_\eta} [k_i - \tfrac{1}{2}(i-1)][a+b+k_i - \tfrac{1}{2}(i-1)] \right\}$$

$$= \frac{(a)_\eta (b)_\eta}{(c)_\eta\, h!} \frac{ab}{C_\eta(I_p)(h+1)} \sum_{i=1}^{\ell} \binom{\eta_i}{\eta} C_{\eta_i}(T_1)$$

$$= \frac{(a)_\eta (b)_\eta}{(c)_\eta\, h!\, C_\eta(I_p)}\, ab\, C_\eta(T_1) \operatorname{Tr}(T_1)$$

where the last line follows from (3.3.11). The theorem now follows from (3.3.16).

Corollary 3.3.1--

The function ${}_2F_1(a,b;c;T,S)$ is the unique solution of (3.3.12) subject to the condition that F may be expressed in the series form

$$F(T,S) = \sum_{k=0}^{\infty} \sum_{\varkappa} \alpha_\varkappa \frac{C_\varkappa(T) C_\varkappa(S)}{C_\varkappa(I_p)}$$

where $F(0,0) = 1$, i.e., $\alpha_{(0)} = 1$.

Proof--

Substitute $F(T,S)$ in (3.3.12) and compare coefficients of $C_\eta(S)$ and then of $C_{\eta_i}(T_1)$ on both sides. This gives the following recurrence relations for the α_h's

$$(h+1)[c+k_i - \tfrac{1}{2}(i-1)]\alpha_{\eta_i} = [a+k_i - \tfrac{1}{2}(i-1)][b+k_i - \tfrac{1}{2}(i-1)]\alpha_\eta .$$

Using the fact that $\alpha_{(0)} = 1$, these recurrence relations determine α_η uniquely as

$$\frac{(a)_\eta (b)_\eta}{(c)_\eta\, h!}$$

Corollary 3.3.2--

The function ${}_1F_0(a;T,S)$ satisfies the partial differential equations

(3.3.20) $$\delta_S F - [a+\tfrac{1}{2}(p-\ell)]\gamma_T F - \eta_T F = \tfrac{1}{2}a(p-1)\mathrm{Tr}(T_1)F$$

and

(3.3.21) $$\delta_S F - \tfrac{1}{2}(p-\ell)\epsilon_S F - a\gamma_T F - \eta_T F = \tfrac{1}{2}a(\ell-1)\mathrm{Tr}(T_1)F .$$

Proof--

(3.3.20) follows from (3.3.12) by putting $c = b = \frac{1}{2}(p-1)$ and (3.3.21) follows from (3.3.12) by putting $c = b = \frac{1}{2}(\ell-1)$.

Theorem 3.3.2--

The function ${}_1F_1(a;c;T,S)$ satisfies the partial differential equation

(3.3.22) $$\delta_S F + [c-\tfrac{1}{2}(p-1)]\epsilon_S F - \gamma_T F = a\ \mathrm{Tr}(T_1)F$$

and the function ${}_0F_1(c;T,S)$ satisfies the partial differential equation

(3.3.23) $$\delta_S F + [c-\tfrac{1}{2}(p-1)]\epsilon_S F = \mathrm{Tr}(T_1)F .$$

Proof--

The theorem can be proved using a proof similar to the proof of Theorem 3.3.1. Alternatively (3.3.22) follows from the differential equation (3.3.12) for ${}_2F_1$ via the confluence

$$\lim_{b\to\infty} {}_2F_1(a,b;c;b^{-1}T,S) = {}_1F_1(a;c;T,S) .$$

Similarly (3.3.23) follows from (3.3.22) and the relation

$$\lim_{a\to\infty} {}_1F_1(a,c;a^{-1}T,S) = {}_0F_1(c;T,S) .$$

Corollary 3.3.3--

The functions ${}_1F_1(a;c;T,S)$ and ${}_0F_1(c;T,S)$ are the unique solutions of (3.3.22) and (3.3.23) subject to the condition that F may be expressed in the form

$$F(T,S) = \sum_{k=0}^{\infty} \sum_{\varkappa} \alpha_{\varkappa} \frac{C_{\varkappa}(T)C_{\varkappa}(S)}{C_{\varkappa}(I_p)}$$

where $F(0,0) = 1$.

Proof--

Follows from Theorem 3.3.2 exactly as Corollary 3.3.1 follows from Theorem 3.3.1.

Corollary 3.3.4--

The function ${}_0F_0(T,S)$ satisfies the partial differential equation

$$\delta_S F - \gamma_T F = \tfrac{1}{2}(p-1)\mathrm{Tr}(T_1)F \ . \tag{3.3.24}$$

Proof--

Put $a = c = \frac{1}{2}(p-1)$ in (3.3.22).

3.4. The Term of Order n^{-1} in the Asymptotic Expansion of ${}_2F_1(\frac{1}{2}n,\frac{1}{2}n;\frac{1}{2}q;P^2,R^2)$.

As indicated in Chapter 2, Section 4, further terms in an asymptotic expansion of ${}_2F_1(n,P^2,R^2)$ could be obtained, at least in principle, by a more detailed analysis of the integrals in Section 3.2. Such an analysis would lead to an asymptotic expansion of the form

$${}_2F_1(n,P^2,R^2) \sim \varphi G$$

where φ is the asymptotic representation for ${}_2F_1(n,P^2,R^2)$ given by Theorem 3.2.1,

$$G = 1 + \frac{P_1}{n} + \frac{P_2}{n^2} + \cdots$$

and the P_i's are functions of the ρ_i^2's and the r_i^2's but do not depend on n .

In this section the partial differential equation satisfied by ${}_2F_1(n, P^2, R^2)$ will be used to compute P_1 .

If S is a $p \times p$ symmetric matrix with latent roots $s_1, \ldots, s_p$, $T = \mathrm{diag}(T_1, 0)$ is a $p \times p$ matrix, and T_1 is a $k \times k$ symmetric matrix with latent roots $t_1, t_2, \ldots, t_k$, then by Theorem 3.3.1 ${}_2F_1(a,b;c;T,S)$ satisfies the following partial differential equation

$$(3.4.1) \quad \sum_{i=1}^{p} s_i \frac{\partial^2 F}{\partial s_i^2} + \sum_{i=1}^{p} \sum_{\substack{j=1 \\ i \neq j}}^{p} \frac{s_i}{s_i - s_j} \frac{\partial F}{\partial s_i} + \{c - \tfrac{1}{2}(p-1)\} \sum_{i=1}^{p} \frac{\partial F}{\partial s_i}$$

$$- (a+b-k+2) \sum_{i=1}^{k} t_i^2 \frac{\partial F}{\partial t_i} - \sum_{i=1}^{k} t_i^3 \frac{\partial^2 F}{\partial t_i^2} - \sum_{i=1}^{k} \sum_{\substack{j=1 \\ i \neq j}}^{k} \frac{t_i^3}{t_i - t_j} \frac{\partial F}{\partial t_i}$$

$$= ab\, F \sum_{i=1}^{k} t_i \; .$$

Let $S = R^2$, $T = P^2$, $a = b = \frac{1}{2}n$, and $c = \frac{1}{2}q$. Then

$$\frac{\partial F}{\partial s_i} = \frac{1}{2r_i} \frac{\partial F}{\partial r_i} \qquad \frac{\partial^2 F}{\partial s_i^2} = \frac{1}{4r_i^2} \frac{\partial^2 F}{\partial r_i^2} - \frac{1}{4r_i^3} \frac{\partial F}{\partial r_i} \qquad i=1, \ldots, p$$

and

$$\frac{\partial F}{\partial t_i} = \frac{1}{2\rho_i} \frac{\partial F}{\partial \rho_i} \qquad \frac{\partial^2 F}{\partial t_i^2} = \frac{1}{4\rho_i^2} \frac{\partial^2 F}{\partial \rho_i^2} - \frac{1}{4\rho_i^3} \frac{\partial F}{\partial \rho_i} \qquad i=1, \ldots, k \; .$$

Substituting the above in (3.4.1) it follows after some simplification that ${}_2F_1(\frac{1}{2}n, \frac{1}{2}n; \frac{1}{2}q; P^2, R^2)$ satisfies

$$(3.4.2)\quad \sum_{i=1}^{p} \frac{\partial^2 F}{\partial r_i^2} + (q-p)\sum_{i=1}^{p} \frac{1}{r_i}\frac{\partial F}{\partial r_i} + 2\sum_{i=1}^{p}\sum_{\substack{j=1\\ i\neq j}}^{p} \frac{r_i}{r_i^2 - r_j^2}\frac{\partial F}{\partial r_i}$$

$$- (2n+3-2k)\sum_{i=1}^{k} \rho_i^3 \frac{\partial F}{\partial \rho_i} - \sum_{i=1}^{k} \rho_i^4 \frac{\partial^2 F}{\partial \rho_i^2}$$

$$- 2\sum_{i=1}^{k}\sum_{\substack{j=1\\ i\neq j}}^{k} \frac{\rho_i^5}{\rho_i^2 - \rho_j^2}\frac{\partial F}{\partial \rho_i} = n^2 F \sum_{i=1}^{k} \rho_i^2 .$$

The partial differential equation satisfied by G is obtained by substituting φG for F in (3.4.2). After a considerable amount of computation it follows that G satisfies

$$(3.4.3)\quad \sum_{i=1}^{p} \frac{\partial^2 G}{\partial r_i^2} + (2n-p-q+1)\sum_{i=1}^{k} \frac{\rho_i}{1 - r_i\rho_i}\frac{\partial G}{\partial r_i} + (q-p)\sum_{i=k+1}^{p} \frac{1}{r_i}\frac{\partial G}{\partial r_i}$$

$$+ 2\sum_{i=k+1}^{p}\sum_{\substack{j=k+1\\ i\neq j}}^{p} \frac{r_i}{r_i^2 - r_j^2}\frac{\partial G}{\partial r_i} - \sum_{i=1}^{k} \rho_i^4 \frac{\partial^2 G}{\partial \rho_i^2}$$

$$+ (q+p-2n-3)\sum_{i=1}^{k} \frac{\rho_i^3}{1 - r_i\rho_i}\frac{\partial G}{\partial \rho_i} + 2\sum_{i=1}^{k} \frac{\rho_i^4 r_i}{1 - r_i\rho_i}\frac{\partial G}{\partial \rho_i}$$

$$= \frac{1}{4}\sum_{i=1}^{k} \rho_i^2 + \frac{1}{2}(p-q)\left(\frac{1}{2}(p-q)+1\right)\sum_{i=1}^{k} r_i^{-2} - \sum_{i=1}^{k}\sum_{\substack{j=1\\ i\neq j}}^{k} \frac{r_i^2}{(r_i^2-r_j^2)^2}$$

$$+ \sum_{i=1}^{k}\sum_{\substack{j=1\\ i\neq j}}^{k} \frac{\rho_i^4\rho_j^2}{(\rho_i^2-\rho_j^2)^2} - \sum_{i=1}^{k}\sum_{j=k+1}^{p} \frac{r_i^2 + r_j^2}{(r_i^2-r_j^2)^2} .$$

The differential equation for P_1 is obtained by substituting

$$1 + \frac{P_1}{n} + \frac{P_2}{n^2} + \cdots$$

for G in (3.4.3) and equating coefficients of powers of n^0. The resulting equation for P_1 is

$$(3.4.4)\qquad 2\sum_{i=1}^{k}\frac{\rho_i}{1-r_i\rho_i}\frac{\partial P_1}{\partial r_i}-2\sum_{i=1}^{k}\frac{\rho_i^3}{1-r_i\rho_i}\frac{\partial P_1}{\partial \rho_i}=\frac{1}{4}\sum_{i=1}^{k}\rho_i^2$$

$$+\frac{1}{4}(p-q)(p-q+2)\sum_{i=1}^{k}r_i^{-2}-\sum_{i=1}^{k}\sum_{\substack{j=1\\ i\neq j}}^{k}\frac{r_i^2}{(r_i^2-r_j^2)^2}$$

$$+\sum_{i=1}^{k}\sum_{\substack{j=1\\ i\neq j}}^{k}\frac{\rho_i^4\rho_j^2}{(\rho_i^2-\rho_j^2)^2}-\sum_{i=1}^{k}\sum_{j=k+1}^{p}\frac{r_i^2+r_j^2}{(r_i^2-r_j^2)^2}\ .$$

Also P_1 satisfies the boundary condition

$$(3.4.5)\qquad P_1(P^2,R^2)=P_1(R^2,P^2)$$

since ${}_2F_1(n,P^2,R^2)={}_2F_1(n,R^2,P^2)$ and $\varphi(n,P^2,R^2)=\varphi(n,R^2,P^2)$.

The general solution of (3.4.4) is

$$P_1=Q+\Psi(u_1,\ldots,u_{p+k-1})$$

where Q is any particular solution, Ψ is an arbitrary function, and the u_i are any $p+k-1$ independent solutions of the system of equations

$$\frac{1-\rho_1 r_1}{2\rho_1}dr_1=\cdots=\frac{1-\rho_k r_k}{2\rho_k}dr_k=\frac{-(1-\rho_1 r_1)}{2\rho_1^3}d\rho_1=\cdots=\frac{-(1-\rho_k r_k)}{2\rho_k^3}d\rho_k\ .$$

Such solutions are easily found to be

$$u_i=\frac{1-r_i\rho_i}{\rho_i}\qquad i=1,\ldots,k$$

$$(3.4.6)\qquad u_{k+i}=\frac{1-r_i\rho_i}{\rho_i^2}-\frac{1-\rho_k r_k}{\rho_k^2}\qquad i=1,\ldots,k-1$$

$$u_{k-1+i}=r_i\qquad i=k+1,\ldots,p\ .$$

We apply the boundary conditions to evaluate Ψ . If we can find a Q which satisfies the boundary conditions then so must Ψ . Examination

of (3.4.6) shows that a function satisfying (3.4.5) must be identically constant. Since there is only one boundary condition for P_1 there is no way to determine the constant.

A particular solution of (3.4.4) is

$$Q = \frac{1}{8}\sum_{i=1}^{k} r_i\rho_i - \frac{1}{8}(q-p)(q-p-2)\sum_{i=1}^{k}(r_i\rho_i)^{-1} + \frac{1}{2}\sum_{i<j}^{k}\sum \frac{(1-r_i\rho_i)(1-r_j\rho_j)(r_i\rho_i+r_j\rho_j)}{(r_i^2-r_j^2)(\rho_i^2-\rho_j^2)} + \frac{1}{2}\sum_{i=1}^{k}\sum_{j=k+1}^{p}\frac{(1-r_i\rho_i)r_i}{(r_i^2-r_j^2)\rho_i} . \tag{3.4.7}$$

Note that $Q(P^2,R^2) = Q(R^2,P^2)$. Hence the general solution of (3.4.4) is

$$P_1 = Q + \eta$$

where Q is given by (3.4.7) and η is a constant.

Additional P_i's could be calculated by the same technique but the partial differential equations for P_i are very complicated when $i \geq 2$.

CHAPTER 4

CANONICAL CORRELATION COEFFICIENTS

4.1. Introduction.

Let $r_1^2 \geq r_2^2 \geq \cdots \geq r_p^2$ be the squared sample canonical correlation coefficients between variates $x_1,\ldots,x_p$ and $y_1,\ldots,y_q$ $(p\leq q)$ calculated from a sample of $N = n+1$ observations from a $(p+q)$-variate normal distribution. The exact joint density function of the r_i^2's is given by (1.3.1) as

$$(4.1.1)\qquad f(R^2) = \prod_{i=1}^{p}(1-\rho_i^2)^{\frac{1}{2}n}\ {}_2F_1(\tfrac{1}{2}n,\tfrac{1}{2}n;\tfrac{1}{2}q;P^2,R^2)$$

$$\times\, C_1 \prod_{i=1}^{p}\{(r_i^2)^{\frac{1}{2}(q-p-1)}(1-r_i^2)^{\frac{1}{2}(n-p-q-1)}\}\prod_{i<j}(r_i^2-r_j^2)\prod_{i=1}^{p}dr_i^2$$

where $1 \geq \rho_1 \geq \rho_2 \geq \cdots \geq \rho_p \geq 0$ are the population canonical correlation coefficients, $R = \mathrm{diag}(r_1,\ldots,r_p)$, $P = \mathrm{diag}(\rho_1,\ldots,\rho_p)$, and

$$C_1 = \frac{\Gamma_p(\frac{1}{2}n)\pi^{\frac{1}{2}p^2}}{\Gamma_p\{\frac{1}{2}(n-q)\}\Gamma_p(\frac{1}{2}q)\Gamma_p(\frac{1}{2}p)}\ .$$

The distribution of the r_i^2's depends only on the ρ_i^2's and that part of the pdf which involves the ρ_i^2's is the marginal likelihood L , i.e.,

$$(4.1.2)\qquad L = \prod_{i=1}^{p}(1-\rho_i^2)^{\frac{1}{2}n}\ {}_2F_1(\tfrac{1}{2}n,\tfrac{1}{2}n;\tfrac{1}{2}q;P^2,R^2)\ .$$

The results of Chapter 3 are used to derive an asymptotic expansion for large n of the pdf (4.1.1) and an asymptotic representation for large n of the likelihood function (4.1.2) under the assumption that the population coefficients satisfy

$$(4.1.3)\qquad 1 > \rho_1 > \rho_2 > \cdots \rho_k > \rho_{k+1} = \cdots = \rho_p = 0\ .$$

These results are then used to study the Bartlett-Lawley tests that the $p - k$ residual population coefficients are zero and to estimate the population coefficients.

4.2. Asymptotic Expansions.

An asymptotic expansion of the joint pdf of $r_1^2, \ldots, r_p^2$ is obtained by substituting the asymptotic expansion of ${}_2F_1(\frac{1}{2}n, \frac{1}{2}n; \frac{1}{2}q; P^2, R^2)$, given by Theorem 3.2.1 and (3.4.7), into (4.1.1). The result is summarized in the following theorem.

Theorem 4.2.1--

An asymptotic expansion for large n of the joint pdf of $r_1^2, \ldots, r_p^2$, the squared sample canonical correlation coefficients, when the population coefficients satisfy (4.13), is

$$\varphi\{1 + \frac{P_1}{n} + O(n^{-2})\} \tag{4.2.1}$$

where

$$\varphi = C_2 \prod_{i=1}^{k} \{(1-r_i\rho_i)^{-n+\frac{1}{2}(p+q-1)}(r_i^2)^{(q-p-2)/4}(1-r_i^2)^{\frac{1}{2}(n-p-q-1)}\} \tag{4.2.2}$$
$$\times \prod_{i<j}^{k} \left\{\frac{r_i^2 - r_j^2}{\rho_i^2 - \rho_j^2}\right\}^{\frac{1}{2}} \prod_{i=1}^{k} \prod_{j=k+1}^{p} (r_i^2 - r_j^2)^{\frac{1}{2}}$$
$$\times \prod_{i=k+1}^{p} \{(r_i^2)^{\frac{1}{2}(q-p-1)}(1-r_i^2)^{\frac{1}{2}(n-q-p-1)}\} \prod_{\substack{k+1 \\ i<j}}^{p} (r_i^2 - r_j^2),$$

$$C_2 = \frac{(\frac{1}{2}n)^{nk-\frac{1}{2}k(p+q)} e^{-nk} \pi^{\frac{1}{2}p^2} \Gamma_k(\frac{1}{2}q)\Gamma_p(\frac{1}{2}n)\Gamma_k(\frac{1}{2}p)}{\{\Gamma_k(\frac{1}{2}n)\}^2 \Gamma_p(\frac{1}{2}p)\Gamma_p(\frac{1}{2}q)\Gamma_p\{\frac{1}{2}(n-q)\}}$$
$$\times \prod_{i=1}^{k} \{(1-\rho_i^2)^{\frac{1}{2}n} \rho_i^{k-\frac{1}{2}(p+q)}\},$$

$$P_1 = \frac{1}{8} \sum_{i=1}^{k} r_i \rho_i - \frac{1}{8}(q-p)(q-p-2) \sum_{i=1}^{k} (r_i \rho_i)^{-1}$$

$$+ \frac{1}{2} \sum\sum_{i<j}^{k} \frac{(1-r_i\rho_i)(1-r_j\rho_j)(r_i\rho_i+r_j\rho_j)}{(r_i^2-r_j^2)(\rho_i^2-\rho_j^2)}$$

$$+ \frac{1}{2} \sum_{i=1}^{k} \sum_{j=k+1}^{p} \frac{(1-r_i\rho_i)r_i}{(r_i^2-r_j^2)\rho_i} + \eta ,$$

η **is a constant, and** $O(n^{-2})$ **are terms which for fixed** ρ_i**'s and** r_i**'s are** $O(n^{-2})$ **as** $n \to \infty$**. Furthermore, for every** $\epsilon > 0$ **let** $B(\epsilon)$ **be the set**

(4.2.3) $B(\epsilon) = \{(r_1^2, \ldots, r_p^2, \rho_1^2, \ldots, \rho_k^2) : r_i^2 - r_{i+1}^2 > \epsilon \quad (0 \leq i \leq p)$

and $\rho_j^2 - \rho_{j+1}^2 > \epsilon \quad (0 \leq j \leq k)$ where $r_0^2 = \rho_0^2 = 1$ and $r_{p+1}^2 = \rho_{k+1}^2 = 0\}$.

Then $f(R^2) \sim \varphi$ uniformly on $B(\epsilon)$ for every $\epsilon > 0$.

The function φ is called the "asymptotic pdf" or "asymptotic distribution" of $r_1^2, \ldots, r_p^2$.

The following corollaries follow easily from Theorem 4.2.1.

Corollary 4.2.1--

An asymptotic representation for large n of the marginal likelihood function L when the ρ_i's satisfy (4.13), is

$$\hat{L} = C_3 \prod_{i=1}^{k} \{(1-\rho_i^2)^{\frac{1}{2}n}(1-\rho_i r_i)^{-n+\frac{1}{2}(p+q-1)} \rho_i^{k-\frac{1}{2}(p+q)}\} \prod_{i<j}^{k} (\rho_i^2-\rho_j^2)^{-\frac{1}{2}}$$

where C_3 is a constant which may depend on n and $r_1^2, \ldots, r_p^2$.
Furthermore, $L \sim \hat{L}$ uniformly on $B(\epsilon)$ for every $\epsilon > 0$, where $B(\epsilon)$ is defined by (4.2.3).

Corollary 4.2.2--

The largest k coefficients $r_1^2, \ldots, r_k^2$ are asymptotically sufficient for the population coefficients $\rho_1^2, \ldots, \rho_k^2$.

Corollary 4.2.3--

The asymptotic conditional pdf of the smallest $p - k$ sample coefficients $r_{k+1}^2, \ldots, r_p^2$ given the largest k coefficients $r_1^2, \ldots, r_k^2$ is

$$(4.2.4)\quad \varphi_C = C_4 \left\{ \prod_{i=1}^{k} \prod_{j=k+1}^{p} (r_i^2 - r_j^2)^{\frac{1}{2}} \right\} \prod_{i=k+1}^{p} \{(r_i^2)^{\frac{1}{2}(q-p-1)} (1-r_i^2)^{\frac{1}{2}(n-q-p-1)}\}$$

$$\times \prod_{\substack{k+1 \\ i<j}}^{p} (r_i^2 - r_j^2) \qquad \text{where } C_4 \text{ is a constant.}$$

Furthermore, if $f(r_{k+1}^2, \ldots, r_p^2 | r_1^2, \ldots, r_k^2)$ is the true conditional distribution of $r_{k+1}^2, \ldots, r_p^2$ given $r_1^2, \ldots, r_k^2$, then $f(r_{k+1}^2, \ldots, r_p^2 | r_1^2, \ldots, r_k^2) \sim \varphi_C$ uniformly on $B(\epsilon)$ for every $\epsilon > 0$, where $B(\epsilon)$ is defined by (4.2.3)

The asymptotic pdf φ, defined by (4.2.2), contains "beta-like" factors linked by factors of the form $r_i^2 - r_j^2$. By making the transformation of variables suggested by Hsu (1941b) it is possible to obtain a "normal-type" approximation which no longer contains linkage factors corresponding to distinct population coefficients. Hsu let

$$(4.2.5)\qquad z_i = \frac{n^{\frac{1}{2}}(r_i^2 - \rho_i^2)}{2\rho_i(1-\rho_i^2)} \qquad (1 \leq i \leq k)$$

and

$$(4.2.6)\qquad z_j = n r_j^2 \qquad (k+1 \leq j \leq p)\ .$$

Making this transformations of variables in (4.2.2) and simplifying gives

Corollary 4.2.4--

The asymptotic joint pdf of $z_1, \ldots, z_p$, when the ρ_i's satisfy (4.1.3), is

$$(4.2.7)\qquad \left\{\prod_{i=1}^{k} g(z_i)\right\} C_5 \exp\left(-\tfrac{1}{2}\sum_{j=k+1}^{p} z_j\right) \prod_{j=k+1}^{p} z_j^{\frac{1}{2}(q-p-1)} \prod_{\substack{k+1\\ i<j}}^{p} (z_i - z_j)$$

$$\times \left[1 + n^{-\frac{1}{2}}\left\{\sum_{i=1}^{k} \frac{z_i(q-p-2+2(p-k)+7\rho_i^2)+(1-3\rho_i^2)z_i^3}{2\rho_i}\right.\right.$$

$$\left.\left. + \sum_{\substack{j=1\\ j\neq i}}^{k} \frac{\rho_i(1-\rho_j^2)z_i}{\rho_i^2-\rho_j^2}\right\} + O(n^{-1})\right]$$

where $g(z_i)$ denotes the standard normal density and

$$C_5 = \frac{\pi^{\frac{1}{2}(p-k)^2}}{2^{\frac{1}{2}(p-k)(q-k)}\Gamma_{p-k}\{\frac{1}{2}(q-k)\}\Gamma_{p-k}\{\frac{1}{2}(p-k)\}}\ .$$

Here $O(n^{-1})$ means terms which are $O(n^{-1})$ as $n \to \infty$ uniformly on any bounded set of z_i's.

The first line of (4.2.7) shows that asymptotically the z_i's which correspond to distinct nonzero ρ_i's are standard normal, independent of z_j $(i \neq j)$. The z_i's corresponding to zero population coefficients are dependent and their asymptotic distribution is the same as the distribution of the latent roots of a $(p-k) \times (p-k)$ matrix having the Wishart distribution $W_{p-k}(q-k, I_{p-k})$.

Hsu (1941b) found the limiting distribution of the z_i's when the population coefficients have arbitrary multiplicity. His result reduces to the first line of (4.2.7) when the nonzero ρ_i's are distinct. Sugiura (1976) found the limiting distribution of the z_i's, up to and including terms of order n^{-1}, when the population coefficients have arbitrary multiplicity and are nonzero. If $p = k$ then (4.2.7) gives the leading term and term of order $n^{-\frac{1}{2}}$ in Sugiura's expansion.

Asymptotic moments (i.e., moments with respect to the asymptotic distribution) of r_i^2 when $\rho_i^2 \neq 0$ is simple, can be obtained from (4.2.7). In particular we have

$$(4.2.8)\quad E(r_i^2) = \rho_i^2 + \frac{(1-\rho_i^2)}{n}\left\{q+p-1-2\rho_i^2+2(1-\rho_i^2)\sum_{\substack{j=1\\j\neq i}}^{k}\frac{\rho_j^2}{\rho_i^2-\rho_j^2}\right\} + O(n^{-2}),$$

$$(4.2.9)\quad E(r_i) = \rho_i + \frac{(1-\rho_i^2)}{2n\rho_i}\left\{q+p-2-\rho_i^2+2(1-\rho_i^2)\sum_{\substack{j=1\\j\neq i}}^{k}\frac{\rho_j^2}{\rho_i^2-\rho_j^2}\right\} + O(n^{-2}),$$

and

$$(4.2.10)\qquad \mathrm{Var}(r_i) = \frac{(1-\rho_i^2)^2}{n} + O(n^{-2}) .$$

These results agree with the results obtained by Lawley (1959).

4.3. Tests of Significance of the Residual Coefficients.

In this section we examine the Bartlett-Lawley tests of the null hypothesis H_k that the residual $p - k$ population canonical correlation coefficients are zero given that $\rho_1^2 > \rho_2^2 > \cdots > \rho_k^2 > 0$. The approach followed here is similar to the one used by James (1969) in connection with the Bartlett-Lawley tests for equality of the $p - k$ smallest latent roots of a covariance matrix.

Bartlett and Lawley studied tests of H_k using test statistics of the form βT_k, where

$$T_k = -\ln \prod_{i=k+1}^{p} (1-r_i^2)$$

and β is a correction factor which may depend on n and $r_1^2,\ldots,r_k^2$, but does not depend on $r_{k+1}^2,\ldots,r_p^2$ (see Chapter 1, Section 3). It is

well known that $n\,T_k$ has an asymptotic χ^2 distribution with $(p-k)(q-k)$ degrees of freedom. The values of β were chosen to improve the rate of convergence of the test statistic βT_k to its asymptotic χ^2 distribution by improving the agreement between the moments of βT_k and the corresponding moments of $\chi^2_{(p-k)(q-k)}$. Lawley (1959) suggested the use of the statistic

$$L_k = \{n - k + \tfrac{1}{2}(p+q+1) + \sum_{i=1}^{k} r_i^{-2}\}T_k .$$

The ρ_i^2's $(1\le i\le k)$ are nuisance parameters in these tests. Since $r_1^2,\ldots,r_k^2$ are asymptotically sufficient for $\rho_1^2,\ldots,\rho_k^2$ (Corollary 4.2.2), the effect of the ρ_i^2's can be eliminated, at least asymptotically, by using the asymptotic conditional distribution of $r_{k+1}^2,\ldots,r_p^2$ given $r_1^2,\ldots,r_k^2$. The technique used here is to find the appropriate factor β by using the conditional density φ_C given by (4.2.4) to compute the mean and variance of T_k. The result confirms Lawley's choice of L_k and provides some information on the accuracy of the approximation.

If $k = 0$ then H_k says that all the ρ_i's are zero. In this case φ_C is the exact null density of $r_1^2,\ldots,r_p^2$ where the linkage factor

$$\prod_{i=1}^{k}\ \prod_{j=k+1}^{p} (r_i^2 - r_j^2)^{\frac{1}{2}}$$

is taken to be unity. Bartlett (1938) proved that in this case

$$-\{n - \tfrac{1}{2}(p+q+1)\}\ln \prod_{j=1}^{p} (1-r_j^2)$$

is asymptotically χ^2 on pq degrees of freedom.

Since $r_i^2 = \rho_i^2 + O_p(n^{-\frac{1}{2}})$ $(1\le i\le k)$ and $r_j^2 = O_p(n^{-1})$ $(k+1\le j\le p)$ (see (4.2.5) - (4.2.7)), r_i^2/r_j^2 is large for large n with probability

arbitrarily close to 1. Thus when testing that the last $p - k$ population coefficients are zero when $k \neq 0$, we could, as a first approximation, ignore the linkage factor

$$\prod_{i=1}^{k} \prod_{j=k+1}^{p} (r_i^2 - r_j^2)^{\frac{1}{2}} .$$

The resulting distribution is approximately the null distribution of the squared sample canonical correlation coefficients with n, p, and q replaced by $n' = n - 2k$, $p' = p - k$, and $q' = q - k$, and suggests the modified statistic

$$-\{n'-\tfrac{1}{2}(p'+q'+1)\} \ln \prod_{j=k+1}^{p} (1-r_j^2) = -\{n-k-\tfrac{1}{2}(p+q+1)\} \ln \prod_{j=k+1}^{p} (1-r_j^2)$$

which is asymptotically $\chi^2_{(p-k)(q-k)}$. The multiplying factor is not quite the same as the one suggested by Bartlett (1947) who replaced n by $n - k$ instead of $n - 2k$. Lawley's (1959) correction factor arises by taking into account the linkage factor.

Let E_C denote expectation with respect to the asymptotic conditional distribution (4.2.4). Then

$$E_C(T_k) = E_C \left\{ -\ln \prod_{i=k+1}^{p} (1-r_i^2) \right\} \tag{4.3.1}$$

$$= E_C \left\{ -\frac{\partial}{\partial h} \left[\prod_{i=k+1}^{p} (1-r_i^2)^h \right]_{h=0} \right\}$$

$$= -\frac{\partial}{\partial h} \left[E_C \left\{ \prod_{i=k+1}^{p} (1-r_i^2)^h \right\} \right]_{h=0}$$

and similarly

$$E_C(T_k^2) = E_C \left\{ \frac{\partial^2}{\partial h^2} \left[\prod_{i=k+1}^{p} (1-r_i^2)^h \right]_{h=0} \right\} \tag{4.3.2}$$

$$= \frac{\partial^2}{\partial h^2}\left[E_C\left\{\prod_{i=k+1}^{p}(1-r_i^2)^h\right\}\right]_{h=0} .$$

We can interchange the order of differentiation and integration in (4.3.1) and (4.3.2) because

$$E_C\left(\sum_{i=k+1}^{p} r_i^2\right)$$

exists and in a neighborhood of $h = 0$

$$\left\{1 - \prod_{i=k+1}^{p}(1-r_i^2)^h\right\} h^{-1} \le 2\sum_{i=k+1}^{p} r_i^2$$

(see for example, Burrill (1972), p. 119). We will actually compute

$$E_C\left\{\prod_{i=k+1}^{p}(1-r_i^2)^h\right\}$$

and then use (4.3.1) and (4.3.2) to calculate $E_C(T_k)$ and $E_C(T_k^2)$.

It follows from Theorem 4.2.1 that for large n the true conditional distribution is

$$(4.3.3) \qquad \varphi_C \times \left\{1 + \frac{1}{2n}\sum_{i=1}^{k}\sum_{j=k+1}^{p}\frac{r_i(1-r_i\rho_i)}{\rho_i(r_i^2-r_j^2)} + O(n^{-2})\right\}$$

where φ_C is the asymptotic conditional pdf defined by (4.2.4). The error in the conditional pdf is $O(n^{-1})$, but since the r_i^2's $(k+1 \le i \le p)$ vary only by $O(n^{-1})$ with probability arbitrarily close to 1, it can be shown by essentially the same argument that we will use to compute

$$E_C\left\{\prod_{i=k+1}^{p}(1-r_i^2)^h\right\},$$

that the relative error in

$$E_C\left\{\prod_{i=k+1}^{p}(1-r_i^2)^h\right\}$$

is $O(n^{-2})$.

Expand the linkage factor in a Taylor series

$$(4.3.4)\qquad \prod_{i=1}^{k}\prod_{j=k+1}^{p}(r_i^2-r_j^2)^{\frac{1}{2}} = \prod_{i=1}^{k}\prod_{j=k+1}^{p}\{r_i^2(1-r_j^2/r_i^2)\}^{\frac{1}{2}}$$

$$= \prod_{i=1}^{k}\prod_{j=k+1}^{p} r_i\left\{1-\frac{r_j^2}{2r_i^2}+O(r_j^4)\right\}$$

$$= \left\{\prod_{i=1}^{k}(r_i)^{p-k}\right\}\left\{1-\tfrac{1}{2}\alpha\sum_{j=k+1}^{p} r_j^2+O(r_j^4)\right\}$$

where

$$\alpha = \sum_{i=1}^{k} r_i^{-2}$$

and $O(r_j^4)$ means terms which are at least fourth order in the r_j $(k+1\leq j\leq p)$. Let

$$(4.3.5)\qquad f_0 = C_6 \prod_{i=k+1}^{p}\{(r_i^2)^{\frac{1}{2}(q-p-1)}(1-r_i^2)^{\frac{1}{2}(n-q-p-1)}\}\prod_{\substack{k+1\\ i<j}}^{p}(r_i^2-r_j^2)$$

where $1 \geq r_{k+1}^2 \geq \cdots \geq r_p^2 \geq 0$ and

$$C_6 = \frac{\pi^{\frac{1}{2}(p-k)^2}\Gamma_{p-k}\{\frac{1}{2}(n-2k)\}}{\Gamma_{p-k}\{\frac{1}{2}(p-k)\}\Gamma_{p-k}\{\frac{1}{2}(n-q-k)\}\Gamma_{p-k}\{\frac{1}{2}(q-k)\}}\ .$$

f_0 is the exact null distribution of the squared canonical correlation coefficients, $r_{k+1}^2,\ldots,r_p^2$, between two sets of variates $w_1,\ldots,w_{p-k}$ and $z_1,\ldots,z_{q-k}$ calculated from a sample of size $n-2k+1$ observations from a $(p+q-2k)$-variate normal distribution. E_N will denote expectation with respect to f_0. Note that f_0 is obtained from φ_c by ignoring the linkage factor, expanding the domain of $r_{k+1}^2,\ldots,r_p^2$ from $r_k^2 \geq r_{k+1}^2 \geq \cdots \geq r_p^2 \geq 0$ to $1 \geq r_{k+1}^2 \geq \cdots \geq r_p^2 \geq 0$, and adjusting the constant. It follows from (4.3.4) and (4.3.5) that

$$(4.3.6)\quad E_C\left\{\prod_{i=k+1}^{p}(1-r_i^2)^h\right\} = C_7\, E_N\left[\prod_{i=k+1}^{p}(1-r_i^2)^h\left\{1-\tfrac{1}{2}\alpha\sum_{j=k+1}^{p} r_j^2 + O(r_j^4)\right\}\right]$$

$$-\int_D \varphi_c \prod_{i=k+1}^{p}(1-r_i^2)^h$$

where $D = \{(r_{k+1}^2,\ldots,r_p^2) : 1 \geq r_{k+1}^2 \geq r_k^2,\ \ r_{k+1}^2 \geq \cdots \geq r_p^2 \geq 0\}$

and

$$C_7 = C_4 C_6^{-1}\prod_{i=1}^{k} r_i^{p-k}\ .$$

We will compute the expectation with respect to the null distribution which occurs in (4.3.6) and then show that the integral in (4.3.6) is of lower order of magnitude. We need the following lemma.

Lemma 4.3.1--

$$\text{(i)}\quad E_N\left\{\left(\sum_{j=k+1}^{p} r_j^2\right)\prod_{i=k+1}^{p}(1-r_i^2)^h\right\} = \frac{(p-k)(q-k)}{n-2k+2h}\,E_0(h)$$

and

$$\text{(ii)}\quad E_N\left\{\left(\sum_{j=k+1}^{p} r_j^2\right)^2\prod_{i=k+1}^{p}(1-r_i^2)^h\right\} = \frac{A\,E_0(h)}{(n+2h-2k-1)(n+2h-2k)(n+2h-2k+2)}$$

where

$$A = (p-k)(q-k)[(n+2h)\{(p-k)(q-k)+2\} + (p-k)(q-k)(1-2k)-2p-2q]\ ,$$

and

$$E_0(h) = E_N\left\{\prod_{i=k+1}^{p}(1-r_i^2)^h\right\}\ .$$

Proof--

The proof uses some properties of multivariate Beta distributions which can be found in Kshirsagar (1972).

Let $u_i = 1 - r_{k+i}^2$ $(1 \leq i \leq p-k)$. The null distribution of $u_1,\ldots,u_{p-k}$ is the same as the distribution of the latent roots of a $(p-k)\times(p-k)$

matrix $U = (u_{ij})$ having a multivariate Beta$\{\frac{1}{2}(n-k-q), \frac{1}{2}(q-k)\}$ distribution.

(i) Note that

$$\sum_{i=k+1}^{p} r_i^2 = \mathrm{tr}(I-U) .$$

Since the diagonal elements of $I - U$ all have the same expectation we need only find the expectation of the first element $\Delta = 1 - u_{11}$ and multiply the result by $(p-k)$.

Put $U = TT'$ where T is a lower triangular $(p-k) \times (p-k)$ matrix. Then $t_{11}^2, \ldots, t_{p-k,p-k}^2$ are independent, t_{ii}^2 has a Beta $\{\frac{1}{2}(n-k-q-i+1), \frac{1}{2}(q-k)\}$ distribution, and $\Delta = 1 - u_{11} = 1 - t_{11}^2$. Hence

$$E_N\left\{\Delta \prod_{i=k+1}^{p} (1-r_i^2)^h\right\} = E_N\left\{(1-t_{11}^2) \prod_{i=1}^{p-k} t_{ii}^{2h}\right\}$$

$$= E_N\left\{\prod_{i=1}^{p-k} t_{ii}^{2h}\right\} - E_N\left\{t_{11}^2 \prod_{i=1}^{p-k} t_{ii}^{2h}\right\} .$$

It is easily shown that

$$E_N\left\{t_{11}^2 \prod_{i=1}^{p-k} t_{ii}^{2h}\right\} = \frac{n - k - q + 2h}{n - 2k + 2h} E_0(h)$$

so that

$$E_N\left\{\Delta \prod_{i=k+1}^{p} (1-r_i^2)^h\right\} = \frac{q - k}{n - 2k + 2h} E_0(h)$$

proving (i).

(ii) Note that

$$\left(\sum_{i=k+1}^{p} r_i^2\right)^2 = \{\mathrm{tr}(I-U)\}^2 .$$

It follows from the fact that the diagonal elements of U are identically distributed that

$$E_N\left\{\left(\sum_{i=k+1}^{p} r_i^2\right)^2 \prod_{i=k+1}^{p}(1-r_i^2)^h\right\} = E_N\left[\left\{\sum_{i=1}^{p-k}(1-u_{ii})\right\}^2 \prod_{i=1}^{p-k} u_{ii}^h\right]$$

$$= E_N[\{(p-k)^2 - 2(p-k)^2 u_{11} + (p-k)u_{11}^2$$

$$+ (p-k)(p-k-1)u_{11}u_{22}\} \prod_{i=1}^{p-k} u_{ii}^h]$$

$$= (p-k)^2 E_N\left\{\prod_{i=1}^{p-k} t_{ii}^{2h}\right\} - 2(p-k)^2 E_N\left\{t_{11}^2 \prod_{i=1}^{p-k} t_{ii}^{2h}\right\}$$

$$+ (p-k)\, E_N\left\{t_{11}^4 \prod_{i=1}^{p-k} t_{ii}^{2h}\right\}$$

$$+ (p-k)(p-k-1)\, E_N\left\{t_{11}^2(t_{21}^2+t_{22}^2) \prod_{i=1}^{p-k} t_{ii}^{2h}\right\} .$$

The first three expectations follow, as in the proof of (i), from the definition of $E_0(h)$ and the fact that the t_{ii}^2's have independent Beta distributions. The results are

$$(4.3.7) \qquad (p-k)^2 E_N\left\{\prod_{i=1}^{p-k} t_{ii}^{2h}\right\} = (p-k)^2 E_0(h) ,$$

$$(4.3.8) \qquad -2(p-k)^2 E_N\left\{t_{11}^2 \prod_{i=1}^{p-k} t_{ii}^{2h}\right\} = \frac{-2(p-k)^2(n+2h-k-q)}{n+2h-2k} E_0(h) ,$$

and

$$(4.3.9) \quad (p-k)\, E_N\left\{t_{11}^4 \prod_{i=1}^{p-k} t_{ii}^{2h}\right\} = \frac{(p-k)(n-k-q+2h+2)(n-k-q+2h)}{(n-2k+2h+2)(n-2k+2h)} E_0(h) .$$

The last expectation is

$$(p-k)(p-k-1)\left[E_N\left\{t_{11}^2 t_{22}^2 \prod_{i=1}^{p-k} t_{ii}^{2h}\right\} + E_N\left\{t_{11}^2 t_{21}^2 \prod_{i=1}^{p-k} t_{ii}^{2h}\right\}\right] .$$

By the same reasoning as in (i) we get

$$(4.3.10)\quad (p-k)(p-k-1)\, E_N\left\{ t_{11}{}^2 t_{22}{}^2 \prod_{i=1}^{p-k} t_{ii}{}^{2h} \right\}$$

$$= \frac{(p-k)(p-k-1)(n-k-q+2h)(n-k-q+2h-1)}{(n-2k+2h)(n-2k+2h-1)}\, E_0(h)\ .$$

To compute $E_N\left\{ t_{11}{}^2 t_{21}{}^2 \prod_{i=1}^{p-k} t_{ii}{}^{2h} \right\}$ partition T as

$$T = \begin{array}{cc} \begin{bmatrix} T_{11} & 0 \\ & \\ T_{21} & T_{22} \end{bmatrix} & \begin{array}{l} 2 \\ \\ p-k-2 \end{array} \\ \begin{array}{cc} 2 & p-k-2 \end{array} & \end{array}$$

where T_{11} and T_{22} are lower triangular and

$$T_{11} = \begin{bmatrix} t_{11} & 0 \\ & \\ t_{21} & t_{22} \end{bmatrix}.$$

T_{11} and T_{22} are independent. $T_{11}T_{11}{}'$ has a multivariate Beta $\{\frac{1}{2}(n-k-q),\ \frac{1}{2}(q-k)\}$ distribution and $T_{22}T_{22}{}'$ has a multivariate Beta $\{\frac{1}{2}(n-k-q-2),\ \frac{1}{2}(q-k)\}$ distribution. It follows that T_{11} , $t_{33}{}^2, \ldots, t_{p-k,p-k}{}^2$ are independent. Put

$$w = \frac{t_{21}}{(1-t_{11}{}^2)^{\frac{1}{2}}(1-t_{22}{}^2)^{\frac{1}{2}}}\ .$$

Then t_{11}, t_{22}, and w are independent, $t_{21}{}^2 = w^2(1-t_{11}{}^2)(1-t_{22}{}^2)$, and $t_{11}{}^2$, $t_{22}{}^2$, and w^2 are distributed as Beta $\{\frac{1}{2}(n-q-k),\ \frac{1}{2}(q-k)\}$, Beta $\{\frac{1}{2}(n-q-k-1),\ \frac{1}{2}(q-k)\}$, and Beta $\{\frac{1}{2}, \frac{1}{2}(q-k-1)\}$, respectively. Therefore

$$E_N\left\{t_{11}^{\ 2}t_{21}^{\ 2}\prod_{i=1}^{p-k}t_{ii}^{\ 2h}\right\} = E_N\left\{t_{11}^{\ 2}(1-t_{11}^{\ 2})(1-t_{22}^{\ 2})w^2\prod_{i=1}^{p-k}t_{ii}^{\ 2h}\right\}$$

$$= E_N(w^2)E_N\left\{t_{11}^{\ 2}(1-t_{11}^{\ 2})(1-t_{22}^{\ 2})\prod_{i=1}^{p-k}t_{ii}^{\ 2h}\right\} .$$

After some simplification we have

$$(4.3.11)\quad (p-k)(p-k-1)E_N\left\{t_{11}^{\ 2}t_{21}^{\ 2}\prod_{i=1}^{p-k}t_{ii}^{\ 2h}\right\}$$

$$= \frac{(n+2h-q-k)(q-k)(p-k)(p-k-1)}{(n+2h-2k+2)(n+2h-2k-1)(n+2h-2k)}E_0(h) .$$

Adding (4.3.7) - (4.3.11) we have

$$E_N\left\{\left(\sum_{j=k+1}^{p}r_j^{\ 2}\right)^2\prod_{i=k+1}^{p}(1-r_i^{\ 2})^h\right\} = \frac{A\ E_0(h)}{(n+2h-2k-1)(n+2h-2k)(n+2h-2k+2)}$$

proving (ii).

It follows from Lemma 4.3.1 and the fact that $0 \le r_j^{\ 2} \le 1$ $(k+1\le j\le p)$ that

$$(4.3.12)\qquad E_N\left[\prod_{i=k+1}^{p}(1-r_i^{\ 2})^h\left\{1-\tfrac{1}{2}\alpha\sum_{j=k+1}^{p}r_j^{\ 2}+O(r_j^4)\right\}\right]$$

$$= E_0(h)\left\{1-\frac{\alpha(p-k)(q-k)}{2(n+2h-2k)}+O(n^{-2})\right\} .$$

We have used the fact that for some constant K

$$E_N\left\{\prod_{i=k+1}^{p}(1-r_i^{\ 2})^h O(r_j^{\ 4})\right\} \le K\,E_N\left\{\prod_{i=k+1}^{p}(1-r_i^{\ 2})^h\left(\sum_{i=k+1}^{p}r_i^{\ 2}\right)^2\right\} = O(n^{-2}) .$$

Now consider the integral in (4.3.6). We have from (4.2.4)

$$\left|\int_D \varphi_c\prod_{i=k+1}^{p}(1-r_i^{\ 2})^h\right| = C_4\int_D\prod_{i=k+1}^{p}\{(r_i^{\ 2})^{\frac{1}{2}(q-p-1)}(1-r_i^{\ 2})^{\frac{1}{2}(n-q-p-1+h}\}$$

$$\times\prod_{\substack{k+1\\ i<j}}^{p}(r_i^{\ 2}-r_j^{\ 2})\prod_{i=1}^{k}\prod_{j=k+1}^{p}(r_i^{\ 2}-r_j^{\ 2})^{\frac{1}{2}}$$

$$\times \prod_{i=k+1}^{p} dr_i^2$$

$$\leq C_4 \int_D (1-r_{k+1}^2)^{\frac{1}{2}(n-q-p-1+h)} \prod_{i=k+1}^{p} dr_i^2 ,$$

since $0 \leq r_i^2 \leq 1$

$$\leq C_4 \int_{r_k^2}^{1} (1-r_{k+1}^2)^{\frac{1}{2}(n-q-p-1+h)} dr_{k+1}^2$$

$$\leq C_4 (1-r_k^2)^{\frac{1}{2}(n-q-p+1+h)}$$

$$= C_4 C_6^{-1} C_6 (1-r_k^2)^{\frac{1}{2}(n-q-p+1+h)}$$

where C_6 is defined by (4.3.5). We can expand C_6 using the following expansion of a gamma function (see, for example, Magnus, et al. (1966), p. 21)

$$\ln \Gamma(n+c) = (n + c - \tfrac{1}{2}) \ln n - n + \tfrac{1}{2} \ln 2\pi + O(n^{-1})$$

to get

$$C_6 = Mn^{\frac{1}{2}(q-k)}\{1 + O(n^{-1})\}$$

where M is a constant which does not depend on n. Then

$$(4.3.13) \quad \left|\int_D \varphi_c \prod_{i=k+1}^{p} (1-r_i^2)^h\right| = C_4 C_6^{-1} Mn^{\frac{1}{2}(q-k)} (1-r_k^2)^{\frac{1}{2}(n-q-p+1+h)}\{1 + O(n^{-1})\}$$

$$= C_7 \, O(\mu^n)$$

for some μ, $1 - r_k^2 < \mu < 1$, where C_7 is defined in (4.3.6). Combining (4.3.6), (4.3.12), and (4.3.13) we have

$$E_C\left\{\prod_{i=k+1}^{p} (1-r_i^2)^h\right\} =$$

$$\frac{C_7 [E_0(h)\{1 - \frac{\alpha(q-k)(p-k)}{2(n+2h-2k)} + O(n^{-2})\} + O(\mu^n)]}{C_7 [E_0(0)\{1 - \frac{\alpha(q-k)(p-k)}{2(n-2k)} + O(n^{-2})\} + O(\mu^n)]}$$

$$= E_0(h)g(h)$$

where

$$g(h) = 1 - \frac{\alpha(p-k)(q-k)h}{(n-2k)(n+2h-2k)} + O(n^{-3}) .$$

From (4.3.1)

$$E_C(T_k) = -\frac{\partial}{\partial h}[E_0(h)g(h)]_{h=0}$$

$$= -E_0{}'(0)g(0) - \frac{\alpha(p-k)(q-k)}{(n-2k)^2} + O(n^{-3})$$

$$= -E_0{}'(0)\{1 + O(n^{-3})\} - \frac{\alpha(p-k)(q-k)}{(n-2k)^2} + O(n^{-3}) .$$

Similarly, from (4.3.2)

$$E_C(T_k{}^2) = \frac{\partial^2}{\partial h^2}[E_0(h)g(h)]_{h=0}$$

$$= E_0{}''(0)\{1 + O(n^{-3})\} + \frac{2\alpha(p-k)(q-k)}{(n-2k)^2}E_0{}'(0) + O(n^{-3}) .$$

Now $-E_0{}'(0) = E_N(T_k)$ and $E_0{}''(0) = E_N(T_k{}^2)$ where T_k is the likelihood ratio statistic for testing the independence of the two sets of variates $w_1,\ldots,w_{p-k}$ and $z_1,\ldots,z_{q-k}$ (see (4.3.5)). Bartlett (1938) proved that

$$E_N(T_k) = \frac{(p-k)(q-k)}{n-2k-\lambda} + O(n^{-3})$$

and

$$E_N(T_k{}^2) = \frac{(p-k)(q-k)\{(p-k)(q-k)+2\}}{(n-2k-\lambda)^2} + O(n^{-3})$$

where

$$\lambda = \tfrac{1}{2}(p+q-2k+1) .$$

Therefore

$$E_C(T_k) = \frac{(p-k)(q-k)}{n-2k-\lambda} - \frac{\alpha(p-k)(q-k)}{(n-2k)^2} + O(n^{-3})$$

$$= \frac{(p-k)(q-k)}{n-2k-\lambda+\alpha} + O(n^{-3})$$

and

$$E_C(T_k^2) = \frac{(p-k)(q-k)\{(p-k)(q-k)+2\}}{(n-2k-\lambda)^2} - \frac{2\alpha(p-k)^2(q-k)^2}{(n-2k)^2(n-2k-\lambda)} + O(n^{-3})$$

$$= \frac{(p-k)(q-k)\{(p-k)(q-k)+2\}}{(n-2k-\lambda+\alpha)^2} + O(n^{-3}) .$$

Thus the appropriate multiplier of T_k is

$$n - 2k - \lambda + \alpha = n - k - \tfrac{1}{2}(p+q+1) + \sum_{i=1}^{k} r_i^{-2} .$$

Theorem 4.3.1 summarizes the previous discussion and Lawley's (1959) result.

Theorem 4.3.1--

The statistic

$$L_k = -\{n - k - \tfrac{1}{2}(p+q+1) + \sum_{i=1}^{k} r_i^{-2}\} \ln \prod_{i=k+1}^{p} (1-r_i^2)$$

is approximately distributed as χ^2 on $(p-k)(q-k)$ degrees of freedom.

If the observed values of $r_1^2,\ldots,r_k^2$ are all near one, then the multiplying factor in L_k is approximately $n - \tfrac{1}{2}(p+q+1)$, which is the value suggested by Bartlett (1938).

4.4. Maximum Likelihood Estimates.

Anderson (1965), James (1966), and Chikuse (1974) have noted, that in the case of the latent roots of a covariance matrix, estimates of population latent roots obtained by maximizing the likelihood based on the sampling distribution of the sample roots, provide a correction for bias.

James (1966) has argued that this is evidence in support of using the marginal distribution of the sample roots and corresponding likelihood function as the basis for inference about the population roots.

Assume that the population coefficients satisfy $1 > \rho_1 > \cdots > \rho_k > \rho_{k+1} = \cdots = \rho_p = 0$. From Corollary 4.2.1 an asymptotic representation for large n of the marginal likelihood function of $\rho_1, \ldots, \rho_k$ is given by

$$\hat{L} = C_3 L_1 L_2$$

where C_3 is a constant,

$$L_1 = \prod_{i=1}^{k} \{(1-\rho_i^2)^{\frac{1}{2}n}(1-\rho_i r_i)^{-n}\}$$

and

$$L_2 = \prod_{i=1}^{k} \{(1-r_i\rho_i)^{\frac{1}{2}(p+q-1)} \rho_i^{k-\frac{1}{2}(p+q)}\} \prod_{i<j}^{k} (\rho_i^2-\rho_j^2)^{-\frac{1}{2}} .$$

The function L_2 is that part of the likelihood which represents the interaction between population coefficients by means of the linkage factors $(\rho_i^2-\rho_j^2)^{-\frac{1}{2}}$. If as a first approximation L_2 is ignored, then the estimate of ρ_i $(i=1,\ldots,k)$ obtained by maximizing L_1 , is the usual maximum likelihood estimate r_i . Using the complete likelihood $\hat{L}$ we have

$$(4.4.1) \quad \ln \hat{L} = \ln C_3 + \tfrac{1}{2}n \sum_{i=1}^{k} \ln(1-\rho_i^2) + \{\tfrac{1}{2}(p+q-1)-n\} \sum_{i=1}^{k} \ln(1-\rho_i r_i)$$
$$+ \{k-\tfrac{1}{2}(p+q)\} \sum_{i=1}^{k} \ln \rho_i - \tfrac{1}{2} \sum_{i<j}^{k}\sum \ln(\rho_i^2-\rho_j^2)$$

and therefore

$$\frac{\partial \ln \hat{L}}{\partial \rho_i} = \frac{-n\rho_i}{1-\rho_i^2} + \{n-\tfrac{1}{2}(p+q-1)\} \frac{r_i}{1-r_i\rho_i} + \frac{(k-\frac{1}{2}(p+q))}{\rho_i} - \sum_{j\neq i}^{k} \frac{\rho_i}{\rho_i^2-\rho_j^2} .$$

Setting

$$\frac{\partial \ln \hat{L}}{\partial \rho_i}$$

equal to zero and solving we obtain the maximum marginal likelihood estimate $\hat{\rho}_i$ of ρ_i as

$$(4.4.2)\quad \hat{\rho}_i = r_i - \frac{(1-r_i^2)}{2nr_i}\left\{ p+q-2+r_i^2+2(1-r_i^2)\sum_{j\neq i}^{k}\frac{r_j^2}{r_i^2-r_j^2}\right\} + O_p(n^{-2}) .$$

It can be shown, using (4.2.5) and (4.2.6), that

$$(4.4.3)\qquad E(\hat{\rho}_i) = \rho_i - \frac{\rho_i(1-\rho_i^2)}{n} + O(n^{-2})$$

and

$$(4.4.4)\qquad \mathrm{Var}(\hat{\rho}_i) = \frac{(1-\rho_i^2)^2}{n} + O(n^{-2}) .$$

These results are similar to those obtained in the case of a simple correlation coefficient. Let $\hat{r}$ be the maximum marginal likelihood estimate based on a sample of size $n+1$ from (x,y), where x and y have a bivariate normal distribution with correlation ρ. Then

$$E(\hat{r}) = \rho - \frac{\rho(1-\rho^2)}{n} + O(n^{-2})$$

and

$$\mathrm{Var}(\hat{r}) = \frac{(1-\rho^2)^2}{n} + O(n^{-2}) .$$

Fisher's z-transform applied to $\hat{r}$ stabilizes the mean and variance. That is, if

$$\xi = \tanh^{-1}\rho = \tfrac{1}{2}\ln\frac{1+\rho}{1-\rho}$$

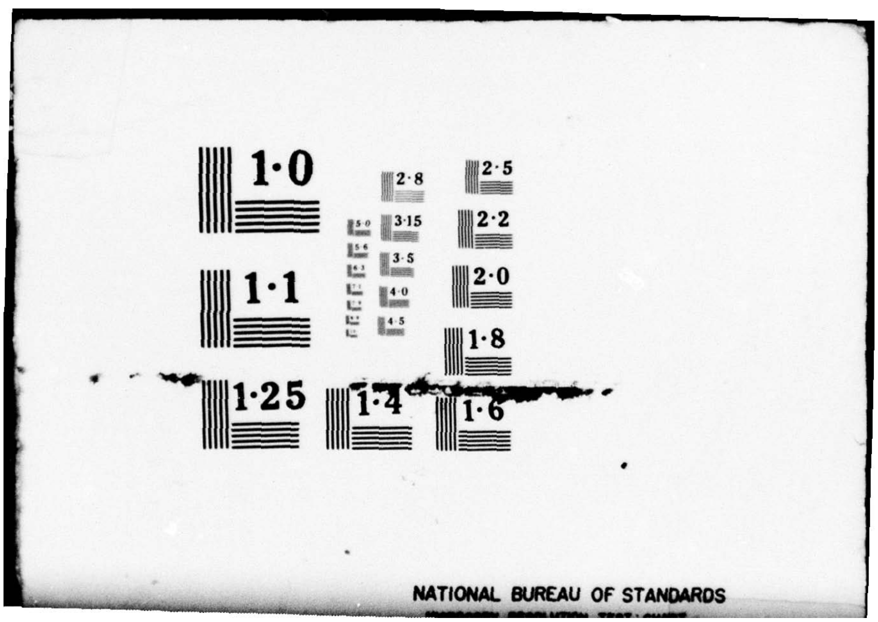
and

$$\hat{\xi} = \tanh^{-1}\hat{r} = \tfrac{1}{2}\ln\frac{1+\hat{r}}{1-\hat{r}} ,$$

AD-A043 772 YALE UNIV NEW HAVEN CONN DEPT OF STATISTICS 12/1
ASYMPTOTIC DISTRIBUTIONS OF LATENT ROOTS IN CANONICAL--ETC(U)
JUN 77 W J GLYNN N00014-75-C-0563
UNCLASSIFIED TR-43 N/L

2 OF 2
ADA
043772

END
DATE
FILMED
9-78
DDC


1·0
2·8
2·5
5·0
5·6
3·15
2·2
3·5
1·1
2·0
4·0
4·5
1·8
1·25
1·4
1·6
NATIONAL BUREAU OF STANDARDS

then

$$E(\hat{\xi}) = \xi + O(n^{-2})$$

and

$$\mathrm{Var}(\hat{\xi}) = n^{-1} + O(n^{-2}) \ .$$

This result suggests applying Fisher's z transformation in the canonical correlation case. Let

$$\xi_i = \tanh^{-1}\rho_i = \tfrac{1}{2}\ln\frac{1+\rho_i}{1-\rho_i} \qquad i=1,\ldots,k$$

and

$$z_i = \tanh^{-1}r_i = \tfrac{1}{2}\ln\frac{1+r_i}{1-r_i} \qquad i=1,\ldots,k$$

where z_i is the usual maximum likelihood estimate of ξ_i. Lawley (1959) notes that this transformation fails to stabilize the mean and variance to any marked extent. In fact z_i has a bias term of order n^{-1}. Let

$$\hat{\xi}_i = \tanh^{-1}\hat{\rho}_i = \tfrac{1}{2}\ln\frac{1+\hat{\rho}_i}{1-\hat{\rho}_i} \qquad i=1,\ldots,k \ .$$

From (4.4.2) we have

$$(4.4.5) \quad \hat{\xi}_i = z_i - \frac{1}{2nr_i}\left\{p+q-2+r_i^2+2(1-r_i^2)\sum_{j\neq i}^{k}\frac{r_j^2}{r_i^2-r_j^2}\right\} + O_p(n^{-2}) \ .$$

Using (3.4.3) and (3.4.4) we can show that

$$E(\hat{\xi}_i) = \xi + O(n^{-2})$$

and

$$\mathrm{Var}(\hat{\xi}_i) = n^{-1} + O(n^{-2}) \ .$$

The effect of Fisher's z-transformation is to provide the same type of bias reduction and variance stabilization as in the correlation coefficient case.

As a numerical example, suppose that $p = q = 3$, $n = 101$ and that the sample values are:

$$r_1 = .630\,, \quad r_2 = .560\,, \quad r_3 = .090$$
$$z_1 = .741\,, \quad z_2 = .633\,, \quad z_3 = .090\,.$$

From (4.4.5) the maximum marginal likelihood estimates, ignoring terms of order n^{-2}, are

$$\hat{\xi}_1 = z_1 - .0682 = .673$$
$$\hat{\xi}_2 = z_2 + .0198 = .653$$
$$\hat{\xi}_3 = z_3 + .0030 = .093\,.$$

An estimate of the standard deviation of $\hat{\xi}_1$ is $s_1 = .1$. The difference between $\hat{\xi}_1$ and z_1 is about $.7s_1$, and is not trivial. It should also be noted that $\hat{\xi}_1 < z_1$ and $\hat{\xi}_3 > z_3$ and hence the effect of L_2 in providing a bias correction is to move the estimates closer together.

One should note that the validity of (4.4.5) depends in large part on the spread of the sample roots. If the sample roots are close together, and this will happen as p gets large, then $\hat{\xi}_i$ will get large due to the terms of order n^{-1} involving $(r_i^2 - r_j^2)^{-1}$.

In the likelihood function of a particular coefficient or set of coefficients the remaining coefficients are nuisance parameters. One can eliminate their effect, at least asymptotically, by deriving the marginal likelihood function from the conditional distribution of the sample coefficients corresponding to those of interest in the population, condi-

tional on the coefficients corresponding to the nuisance parameters as in Section 2. An alternative approach, suggested by James (1966) in the case of the latent roots of a covariance matrix, is to average the joint likelihood function given in Corollary 4.2.1 over some reasonable posterior distribution of the nuisance parameters. James argued that if these nuisance parameters are somewhat removed from the parameters of interest then the likelihood obtained by this averaging would not be very different from that obtained from the joint likelihood by replacing the nuisance parameters by their maximum likelihood estimates. The log-likelihood of a single population coefficient ρ_i would be, from (4.4.1),

$$K + \tfrac{1}{2}n \ln(1-\rho_i^2) + \{\tfrac{1}{2}(p+q-1)-n\} \ln(1-r_i\rho_i) + \{k-\tfrac{1}{2}(p+q)\} \ln \rho_i$$

$$- \tfrac{1}{2} \sum_{j=1}^{i-1} \ln\left(1 - \frac{\rho_i^2}{\hat{\rho}_j^2}\right) - \tfrac{1}{2} \sum_{j=i+1}^{k} \ln\left(\frac{\rho_i^2}{\hat{\rho}_j^2} - 1\right)$$

where K does not depend on ρ_i. The log likelihood of a group of population coefficients would be obtained from (4.4.1) in a similar manner.

CHAPTER 5

AN ASYMPTOTIC EXPANSION OF ${}_1F_1\{\frac{1}{2}(n_1+n_2);\frac{1}{2}n_1;\frac{1}{2}n_2\Theta,L\}$ FOR LARGE n_2

5.1. Introduction.

Let ${}_1F_1(n_2,\Theta,L)$ denote the hypergeometric function ${}_1F_1\{\frac{1}{2}(n_1+n_2);\frac{1}{2}n_1;\frac{1}{2}n_2\Theta,L\}$ where $L = \mathrm{diag}(\ell_1,\ldots,\ell_p)$ with $1 > \ell_1 > \ell_2 > \cdots > \ell_p > 0$ and $\Theta = \mathrm{diag}(\theta_1,\ldots,\theta_p)$ with $\infty > \theta_1 > \cdots > \theta_p > 0$. This hypergeometric function occurs in the pdf of latent roots which arise in MANOVA and discriminant analysis (see Chapter 1, Section 4). Assume that $\Theta = \mathrm{diag}(\Theta_1,0)$ where $\Theta_1 = \mathrm{diag}(\theta_1,\ldots,\theta_k)$ $(p \geq k)$.

The technique presented in Chapter 2 is used to derive an asymptotic expansion of ${}_1F_1(n_2,\Theta,L)$ for large n_2.

5.2. An Asymptotic Representation of ${}_1F_1\{\frac{1}{2}(n_1+n_2);\frac{1}{2}n_1;\frac{1}{2}n_2\Theta,L\}$ for Large n_2.

An asymptotic representation of ${}_1F_1(n_2,\Theta,L)$ will be obtained using Theorem 2.3.1. The derivation is essentially the same as the derivation of the asymptotic representation of ${}_2F_1(n,P^2,R^2)$ for large n which was presented in Chapter 3, Section 2. The reader is referred to that section for the details of this derivation.

We begin by applying the results of Chapter 1, Section 2 to express ${}_1F_1(n_2,\Theta,L)$ as a multiple integral. From (1.2.6) we have

$${}_1F_1(n_2,\Theta,L) = \int_{O(p)} {}_1F_1^{(p)}\{\tfrac{1}{2}(n_1+n_2);\tfrac{1}{2}n_1;\tfrac{1}{2}n_2\Theta^{\frac{1}{2}}H'LH\Theta^{\frac{1}{2}}\}\,(dH)\,.$$

Partition H into two submatrices H_1 and H_2 consisting of the first k and last $p-k$ columns, respectively. It follows from (1.2.5) and

the fact that $\Theta = \operatorname{diag}(\Theta_1, 0)$ that

$$ {}_1F_1(n_2, \Theta, L) = \int_{O(p)} {}_1F_1^{(k)}\{\tfrac{1}{2}(n_1+n_2); \tfrac{1}{2}n_1; \tfrac{1}{2}n_2 \Theta_1^{\frac{1}{2}} H_1' L H_1 \Theta_1^{\frac{1}{2}}\} (dH) . $$

Because the integrand does not depend on H_2, we can integrate over H_2 using Lemma 3.2.1. The result is

$$ {}_1F_1(n_2, \Theta, L) = \int_{V(k,p)} {}_1F_1^{(k)}\{\tfrac{1}{2}(n_1+n_2); \tfrac{1}{2}n_1; \tfrac{1}{2}n_2 \Theta_1^{\frac{1}{2}} H_1' L H_1 \Theta_1^{\frac{1}{2}}\} (dH_1) . $$

For $n_1 + n_2 > k - 1$ we may apply the Laplace transform relation (1.2.3) to obtain

$$ {}_1F_1(n_2, \Theta, L) = [\Gamma_k\{\tfrac{1}{2}(n_1+n_2)\}]^{-1} \int_{V(k,p)} \int_{X>0} \operatorname{etr}(-X)(\det X)^{\frac{1}{2}(n_1+n_2-k-1)} $$

$$ \times {}_0F_1\{\tfrac{1}{2}n_1; \tfrac{1}{2}n_2 X^{\frac{1}{2}} \Theta_1^{\frac{1}{2}} H_1' L H_1 \Theta_1^{\frac{1}{2}} X^{\frac{1}{2}}\} (dX)(dH_1) $$

where X is a $k \times k$ positive definite matrix. It follows by applying Bessel's integral (1.2.7) to the ${}_0F_1$ function that

$$ {}_1F_1(n_2, \Theta, L) = [\Gamma_k\{\tfrac{1}{2}(n_1+n_2)\}]^{-1} \int_{V(k,p)} \int_{X>0} \int_{O(n_1)} \operatorname{etr}(-X)(\det X)^{\frac{1}{2}(n_1+n_2-k-1)} $$

$$ \times \operatorname{etr}\{2^{\frac{1}{2}} n_2^{\frac{1}{2}} [X^{\frac{1}{2}} \Theta_1^{\frac{1}{2}} H_1' L^{\frac{1}{2}} : 0] M_1\} (dM)(dX)(dH_1) $$

where M_1 is the $n_1 \times k$ matrix formed by the first k columns of $M \in O(n_1)$ and 0 in $[X^{\frac{1}{2}} \Theta_1^{\frac{1}{2}} H_1' L^{\frac{1}{2}} : 0]$ is the $k \times (n_1-k)$ zero matrix. Since the integrand does not depend on the last $n_1 - k$ columns of M, we may integrate over them using Lemma 3.2.1 to obtain

$$ {}_1F_1(n_2, \Theta, L) = [\Gamma_k\{\tfrac{1}{2}(n_1+n_2)\}]^{-1} \int_{V(k,p)} \int_{X>0} \int_{V(k,n_1)} \operatorname{etr}(-X)(\det X)^{\frac{1}{2}(n_1+n_2-k-1)} $$

$$ \times \operatorname{etr}\{2^{\frac{1}{2}} n_2^{\frac{1}{2}} [X^{\frac{1}{2}} \Theta_1^{\frac{1}{2}} H_1' L^{\frac{1}{2}} : 0] M_1\} (dM_1)(dX)(dH_1) . $$

Let $X = \tfrac{1}{2}n_2 G' V^2 G$ where $V = \operatorname{diag}(v_1, \ldots, v_k)$ with $v_1 > v_2 > \cdots v_k > 0$, and G is a $k \times k$ orthogonal matrix with positive elements in the first

column. The Jacobian of this transformation, $J\{X \to (V,G)\}$ is given by

$$J\{X \to (V,G)\} = \frac{(\frac{1}{2}n_2)^{\frac{1}{2}k(k+1)} \pi^{\frac{1}{2}k^2} 2^{2k}}{\Gamma_k(\frac{1}{2}k)} \det V \prod_{i<j}^{k} (v_i^2 - v_j^2)$$

(see (3.2.3) and the subsequent discussion). Using the fact that the resulting integrand is invariant under transformations which change the sign of the elements in one row of G, we have

$$(5.2.1) \qquad {}_1F_1(n_2, \Theta, L) = C_{n_2} \int_{\Lambda} \eta(y)\{g(y, \theta, \ell)\}^{n_2} dy$$

where

$$C_{n_2} = \frac{(\frac{1}{2}n_2)^{\frac{1}{2}k(n_1+n_2)} \pi^{\frac{1}{2}k^2} 2^k}{\Gamma_k\{\frac{1}{2}(n_1+n_2)\}\Gamma_k(\frac{1}{2}k)} ,$$

$$\eta(y) = (\det V)^{n_1-k} \prod_{i<j}^{k} (v_i^2 - v_j^2) ,$$

$$g(y, \theta, \ell) = \operatorname{etr}\{-\tfrac{1}{2}V^2 + [G'VG\Theta_1^{\frac{1}{2}}H_1'L^{\frac{1}{2}}:0]M_1\} \det V ,$$

$$\Lambda = V(k,p) \times O(k) \times D_V \times V(k,n_1) ,$$

$$D_V = \{(v_1, \ldots, v_k) : v_1 > v_2 > \cdots > v_k > 0\} ,$$

$$\theta = (\theta_1, \ldots$$

$$\ell = (\ell_1, \ldots, \ell_p) ,$$

and y is a point in Λ. Finally make the transformations $G \to G$, $H_1 \to H_1$, and $M_1G' \to E_1$ in (5.2.1) to get

$$(5.2.2) \qquad {}_1F_1(n_2, \Theta, L) = C_{n_2} \int_{\Lambda} \eta(y) \{h(y, \theta, \ell)\}^{n_2} dy$$

where

$$h(y,\theta,\ell) = \operatorname{etr}\{-\tfrac{1}{2}V^2 + [VG\Theta_1^{\frac{1}{2}}H_1{}'L^{\frac{1}{2}}:0]E_1\}\det V \ .$$

We now determine the maximum value of h for fixed Θ and L and show that the maximum is obtained at a finite number of interior points of Λ. The result is contained in the following lemma.

Lemma 5.2.1--

Let

(i) $\Theta_1 = \operatorname{diag}(\theta_1,\ldots,\theta_k)$ with $\theta_1 > \theta_2 > \cdots \theta_k > 0$;

(ii) $L = \operatorname{diag}(\ell_1,\ldots,\ell_p)$ with $\ell_1 > \ell_2 > \cdots > \ell_p > 0$ $(p \geq k)$;

(iii) $V = \operatorname{diag}(v_1,\ldots,v_k)$ with $v_1 > v_2 > \cdots > v_k > 0$;

(iv) $G = (g_{ij}) \in O(k)$, $(1 \leq i, j \leq k)$;

(v) $E_1 = (e_{ij}) \in V(k,n_1)$ $(1 \leq i \leq n_1, 1 \leq j \leq k)$;

(vi) $H_1 = (h_{ij}) \in V(k,p)$ $(1 \leq i \leq p, 1 \leq j \leq k)$; and

(vii) $h = \operatorname{etr}\{-\tfrac{1}{2}V^2 + [VG\Theta_1^{\frac{1}{2}}H_1{}'L^{\frac{1}{2}}:0]E_1\}\det V$.

Then the maximum value of h for fixed Θ_1 and L is

$$2^{-k}e^{-\frac{1}{2}k}\prod_{i=1}^{k}\{(\ell_i\theta_i)^{\frac{1}{2}} + (\ell_i\theta_i+4)^{\frac{1}{2}}\}\exp\left[\frac{1}{4}\sum_{i=1}^{k}\{\ell_i\theta_i + (\ell_i\theta_i)^{\frac{1}{2}}(\ell_i\theta_i+4)^{\frac{1}{2}}\}\right]$$

and the maximum is obtained if and only if

$$v_i = \tfrac{1}{2}\{(\ell_i\theta_i)^{\frac{1}{2}} + (\ell_i\theta_i+4)^{\frac{1}{2}}\} \quad (1 \leq i \leq k) , \quad G = \operatorname{diag}(\pm 1,\ldots,\pm 1) ,$$

$$H_1 = \begin{bmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \\ \cdot & \cdot & \cdot \\ & 0 & \end{bmatrix} \quad \text{and} \quad E_1 = \begin{bmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \\ \cdot & \cdot & \cdot \\ & 0 & \end{bmatrix}$$

where G , H_1 , and E_1 satisfy the following constraints

$$g_{ii}h_{ii}e_{ii} = 1 \qquad (1 \leq i \leq k) .$$

Proof--

Let $h_1 = e^T$ where $T = \operatorname{tr}\{[VG\Theta_1^{\frac{1}{2}}H_1{}'L^{\frac{1}{2}}:0]E_1\}$ and $h_2 = \operatorname{etr}(-\tfrac{1}{2}V^2)\det V$. Then $h = h_1h_2$. Maximizing T is equivalent to

maximizing h_1 since h_1 is a strictly increasing function of T. It follows from Corollary 1 of the Appendix, that for fixed V, $\Theta_1^{\frac{1}{2}}$, and $L^{\frac{1}{2}}$, the maximum value of T is

$$\sum_{i=1}^{k} v_i \theta_i^{\frac{1}{2}} \ell_i^{\frac{1}{2}}$$

and the maximum is obtained if and only if $G = \operatorname{diag}(\pm 1, \ldots, \pm 1)$,

$$H_1 = \begin{bmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \\ \cdot & \cdot & \cdot \\ & 0 & \end{bmatrix} \quad \text{and} \quad E_1 = \begin{bmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \\ \cdot & \cdot & \cdot \\ & 0 & \end{bmatrix}$$

where G, H_1, and E_1 satisfy the constraints $g_{ii}h_{ii}e_{ii} = 1$ $(1 \leq i \leq k)$. Since the location of the maxima does not depend on V, we may maximize h in two stages as follows.

$$\max_{V,G,H_1,E_1} h = \max_{V}\{h_2(\max_{G,H_1,E_1} h_1)\} .$$

The problem of maximizing h thus reduces to maximizing

$$g_1(V) = \exp\left\{\sum_{i=1}^{k} (-\tfrac{1}{2}v_i^2 + v_i\theta_i^{\frac{1}{2}}\ell_i^{\frac{1}{2}} + \ln v_i)\right\}$$

$$= \prod_{i=1}^{k} \{\exp(-\tfrac{1}{2}v_i^2 + v_i\theta_i^{\frac{1}{2}}\ell_i^{\frac{1}{2}} + \ln v_i)\} .$$

Consider the function $g_2(x)$ defined for all $x > 0$ by

$$g_2(x) = -\tfrac{1}{2}x^2 + x\alpha + \ln x$$

with $\alpha > 0$. It is easily shown that g_2 has an absolute maximum value of

$$-\tfrac{1}{2} - \ln 2 + \ln\{\alpha + (\alpha^2+4)^{\frac{1}{2}}\} + \frac{\alpha}{4}\{\alpha + (\alpha^2+4)^{\frac{1}{2}}\}$$

which is obtained if and only if $x = \frac{1}{2}\{\alpha + (\alpha^2 + 4)^{\frac{1}{2}}\}$. The lemma follows from this result and the fact that

$$\tfrac{1}{2}\{(\ell_1\theta_1)^{\frac{1}{2}} + (\ell_1\theta_1+4)^{\frac{1}{2}}\} > \cdots > \tfrac{1}{2}\{(\ell_k\theta_k)^{\frac{1}{2}} + (\ell_k\theta_k+4)^{\frac{1}{2}}\} > 0 .$$

By the same argument that led to (3.2.6), h obtains its maximum at 2^{2k} interior points of Λ and Λ can be partitioned into 2^{2k} disjoint sets Λ_i $(1 \leq i \leq 2^{2k})$ such that each Λ_i contains exactly one of these points in its interior. Furthermore, Λ_i is of the form

$$\Lambda_i = M_{i1} \times M_{i2} \times D_V \times M_{i3} \tag{5.2.3}$$

where $M_{i1} \subset V(k,p)$, $M_{i2} \subset O(k)$, and $M_{i3} \subset V(k,n_1)$.

Now write (5.2.2) as the sum of 2^{2k} integrals where each integral is the integral of $C_{n_2} \psi h^{n_2}$ over one of the Λ_i's. The only part of the integrand which depends on G, H_1, or E_1 is λ, where $\lambda = \operatorname{etr}(n_2[VG\,\Theta_1^{\frac{1}{2}}H_1'L^{\frac{1}{2}}:0]E_1)$. λ is invariant under transformations of the form

$$G \to GS_1, \qquad H_1 \to \operatorname{diag}(S_2, I_{p-k})H_1S_1, \qquad \text{and}$$

$$E_1 \to \operatorname{diag}(S_2, I_{n_1-k})E_1$$

where S_1 and S_2 are $k \times k$ diagonal matrices with ± 1 on the diagonal. The integral over Λ_i can be transformed by an appropriate choice of S_1 and S_2 to an integral of the form

$$J_i = C_{n_2} \int_{\Omega_i} \psi(y)[h(y,\theta,\ell)]^{n_2} dy\,. \tag{5.2.4}$$

S_1 and S_2 are chosen so that h obtains its maximum at the interior point $\beta(\theta,\ell)$ of Ω_i defined by

$$G = I_k\,, \quad H_1 = \begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix}, \quad E_1 = \begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix}, \quad \text{and} \tag{5.2.5}$$

$$v_i = \tfrac{1}{2}\left((\ell_i\theta_i)^{\frac{1}{2}} + (\ell_i\theta_i + 4)^{\frac{1}{2}}\right) \qquad (1 \leq i \leq k)\,.$$

It follows from (5.2.3) that Ω_i is of the form

(5.2.6) $$\Omega_i = N_{i1} \times N_{i2} \times D_V \times N_{i3}$$

where $N_{i1} \subset V(k,p)$, $N_{i2} \subset O(k)$ and $N_{i3} \subset V(k,n_1)$.

Combining these results we have

(5.2.7) $${}_1F_1(n_2,\Theta,L) = \sum_{i=1}^{2^{2k}} J_i \ .$$

We will derive an asymptotic representation φ_i of J_i . In fact $\varphi_i = \varphi$ for all i , and hence by (5.2.7), an asymptotic representation for ${}_1F_1(n_2,\Theta,L)$ is given by $2^{2k}\varphi$.

We can show, by the same argument that led to (3.2.11), that for each i $(1 \le i \le 2^{2k})$ Ω_i can be partitioned as

$$\Omega_i = \Xi \cup (\Omega_i - \Xi) \ .$$

All orthogonal matrices in Ξ are proper, $\beta(\theta,\ell)$ is an interior point of Ξ for all (θ,ℓ) , and

(5.2.8) $$\Xi = N_1 \times N_2 \times D_V \times N_3$$

where N_1 , N_2 , and N_3 are neighborhoods of

$$\begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix} \in V(k,p) \ , \quad I_k \ , \quad \text{and} \quad \begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix} \in V(k,n_1) \ , \quad \text{respectively.}$$

J_i may be written as

(5.2.9) $$J_i = C_{n_2}(L_1 + L_{i2})$$

where

(5.2.10) $$L_1 = \int_{\Xi} \Pi(y)\{h(y,\theta,\ell)\}^{n_2} dy$$

and

(5.2.11) $$L_{i2} = \int_{\Omega_i - \Xi} \Pi(y)\{h(y,\theta,\ell)\}^{n_2} dy \ .$$

L_1 and L_{12} will be treated separately. We will use Theorem 2.3.1 to derive an asymptotic representation of L_1 , and then we will prove that L_{12} is asymptotically of lower order of magnitude than L_1 .

Using the notation of Theorem 2.3.1 let

(5.2.12) $A = \{(\theta,\ell): \infty > \theta_1 > \theta_2 > \cdots > \theta_k > 0 \ , \quad 1 > \ell_1 > \ell_2 > \cdots > \ell_p > 0\}$

and for each ϵ , $0 < \epsilon < \frac{1}{2}$, and $K > 0$, let $B = B(K,\epsilon)$ be the set

(5.2.13) $B = \{(\theta,\ell) \in A : K > \theta_1 \ , \quad \theta_i - \theta_{i+1} > \epsilon \quad (1 \le i \le k)$ and

$\ell_j - \ell_{j+1} > \epsilon \quad (1 \le j \le p)$, where $\theta_{k+1} = \ell_{p+1} = 0\}$.

Derivation of an Asymptotic Representation for L_1 .

G , H_1 , and E_1 can be parameterized in Ξ by

$$G = \exp(S) \ ,$$

$$[H_1 : -] = \exp(W) = \exp\left(\begin{bmatrix} W_{11} & W_{12} \\ -W_{12}' & 0 \end{bmatrix}\right) \ ,$$

and

$$[E_1 : -] = \exp(Z) = \exp\left(\begin{bmatrix} Z_{11} & -Z_{21}' \\ Z_{21} & 0 \end{bmatrix}\right)$$

where S , W_{11} , and Z_{11} are $k \times k$ skew-symmetric matrices, W_{12} is $k \times (p-k)$, and Z_{21} is $(n_1-k) \times k$. The Jacobians of these transformations are

$$J(G \to S) = \frac{\Gamma_k(\frac{1}{2}k)}{2^k \pi^{\frac{1}{2}k^2}} \{1 + O(s_{ij}^{\ 2})\} \ ,$$

$$J\{H_1 \to (W_{11}, W_{12})\} = \frac{\Gamma_k(\frac{1}{2}p)}{2^k \pi^{\frac{1}{2}kp}} \{1 + O(w_{ij}^{\ 2})\} \ ,$$

and

$$J\{E_1 \to (Z_{11}, Z_{21})\} = \frac{\Gamma_k(\frac{1}{2}n_1)}{2^k \pi^{\frac{1}{2}n_1 k}} \{1 + O(z_{ij}^{\ 2})\}$$

(see (3.2.17) - (3.2.20)). In this section the notation $O(s_{ij}^{m})$ **means** **terms** in the s_{ij} which are at least of order m .

The point $\beta(\theta, \ell)$ defined by (5.2.5) is mapped into the point $\xi(\theta, \ell)$ defined by

(5.2.14) $$S = W_{11} = Z_{11} = 0\ ,\quad W_{12} = 0\ ,\quad Z_{21} = 0\ ,\qquad \text{and}$$

$$v_i = \tfrac{1}{2}\{(\ell_i\theta_i)^{\frac{1}{2}} + (\ell_i\theta_i+4)^{\frac{1}{2}}\} \qquad (1 \le i \le k)$$

and Ξ is mapped into a set D which contains $\xi(\theta, \ell)$ in its interior. It follows after some simplification that

$$\mathrm{tr}\{[VG\,\Theta_1^{\frac{1}{2}}H_1{}'L^{\frac{1}{2}}:0]E_1\} = \sum_{i=1}^{k} v_i\theta_i^{\frac{1}{2}}\ell_i^{\frac{1}{2}} + \Psi$$

where

(5.2.15) $$\Psi = -\sum_{i<j}^{k}\{\tfrac{1}{2}(v_i\theta_i^{\frac{1}{2}}\ell_i^{\frac{1}{2}} + v_j\theta_j^{\frac{1}{2}}\ell_j^{\frac{1}{2}})(s_{ij}^{\;2} + w_{ij}^{\;2} + z_{ij}^{\;2})$$

$$+ (v_i\theta_j^{\frac{1}{2}}\ell_i^{\frac{1}{2}} + v_j\theta_i^{\frac{1}{2}}\ell_j^{\frac{1}{2}})s_{ij}w_{ij} + (v_i\theta_j^{\frac{1}{2}}\ell_j^{\frac{1}{2}} + v_j\theta_i^{\frac{1}{2}}\ell_i^{\frac{1}{2}})s_{ij}z_{ij}$$

$$+ (v_i\theta_i^{\frac{1}{2}}\ell_j^{\frac{1}{2}} + v_j\theta_j^{\frac{1}{2}}\ell_i^{\frac{1}{2}})w_{ij}z_{ij}\}$$

$$- \sum_{i=1}^{k}\sum_{j=k+1}^{p}[\tfrac{1}{2}v_i\theta_i^{\frac{1}{2}}\ell_i^{\frac{1}{2}}\{z_{ij}^{\;2} + w_{ij}^{\;2}\} + v_i\theta_i^{\frac{1}{2}}\ell_j^{\frac{1}{2}}w_{ij}z_{ij}]$$

$$- \tfrac{1}{2}\sum_{i=1}^{k}\sum_{j=p+1}^{n_1} v_i\theta_i^{\frac{1}{2}}\ell_i^{\frac{1}{2}}z_{ij}^{\;2}\ .$$

Combining these results we have

(5.2.16) $$L_1 = \int_D \tau(x)\{f(x, \theta, \ell)\}^{n_2}dx$$

where

$$f(x, \theta, \ell) = \exp\left\{\sum_{i=1}^{k}(-\tfrac{1}{2}v_i^{\;2} + \theta_i^{\frac{1}{2}}\ell_i^{\frac{1}{2}}v_i) + \Psi\right\}\prod_{i=1}^{k} v_i\ ,$$

$$\tau(x) = \frac{\Gamma_k(\frac{1}{2}k)\Gamma_k(\frac{1}{2}p)\Gamma_k(\frac{1}{2}n_1)}{2^{3k}\pi^{\frac{1}{2}k(k+p+n_1)}} \prod_{i=1}^{k} v_i^{n_1-k} \prod_{i<j}^{k} (v_i^2 - v_j^2)$$

$$\times J(G\to S)J\{H_1\to(W_{11},W_{12})\}J\{E_1\to(Z_{11},Z_{21})\} ,$$

D is the image of Ξ, and $x \in D$.

An asymptotic representation of L_1 will now be obtained by applying Theorem 2.3.1 to (5.2.16). We must verify that the conditions of the theorem are satisfied. A basic difference between the functions ${}_2F_1(n,P^2,R^2)$ considered in Chapter 3 and ${}_1F_1(n_2,\Theta,L)$, is that the population and sample roots are bounded in the former but only the ℓ_i's are bounded in the latter. It turns out that the conditions of Theorem 2.3.1 are satisfied by the integral representation (5.2.16) for ${}_1F_1(n_2,\Theta,L)$ only on bounded sets B of (θ,ℓ)'s. For this reason we had to introduce the parameter K in the definition (5.2.13) of $B = B(K,\epsilon)$. K is an upper bound for the θ_i's.

Verification of the conditions of Theorem 2.3.1.

Conditions (i)-(ix) can be verified by essentially the same arguments that were used in Chapter 3 to verify the corresponding conditions for the integral L_1 defined by (3.2.23). In effect we need only verify condition (v) and prove that Lemma 3.2.3, which is used to verify condition (vii), holds for $h(y,\theta,\ell)$ defined by (5.2.2).

(v) We have to show that

$$0 < \inf_B \gamma_k^2(\theta,\ell) \quad \text{and} \quad \sup_B \gamma_1^2(\theta,\ell) < \infty$$

where $\gamma_1^2(\theta,\ell) \geq \gamma_2^2(\theta,\ell) \geq \cdots \geq \gamma_k^2(\theta,\ell)$ are the latent roots of $\Omega\{\xi(\theta,\ell),\theta,\ell\}$. $\Omega\{x,\theta,\ell\} = (w_{ij}(x,\theta,\ell))$ is the $k \times k$ matrix defined for all $x \in D$ and $(\theta,\ell) \in A$ by

$$w_{ij}(x, \theta, \ell) = - \frac{\partial^2 \ln f}{\partial x_i \partial x_j}(x, \theta, \ell) .$$

To compute $\Omega\{\xi(\theta, \ell), \theta, \ell\}$ we need the second partial derivatives of $\psi(x, \theta, \ell) = -\ln f(x, \theta, \ell)$ evaluated at $\{\xi(\theta, \ell), \theta, \ell\}$. It follows from (5.2.15), that for fixed (θ, ℓ) the only non-zero second partial derivatives of ψ evaluated at $\{\xi(\theta, \ell), \theta, \ell\}$ are

$$\frac{\partial^2 \psi}{\partial v_i^2} = \frac{2(\theta_i \ell_i + 4)^{\frac{1}{2}}}{(\theta_i \ell_i)^{\frac{1}{2}} + (\theta_i \ell_i + 4)^{\frac{1}{2}}} \qquad (1 \le i \le k)$$

$$\frac{\partial^2 \psi}{\partial s_{ij}^2} = \frac{\partial^2 \psi}{\partial w_{ij}^2} = \frac{\partial^2 \psi}{\partial z_{ij}^2} = \tfrac{1}{2}(\ell_i \theta_i)^{\frac{1}{2}}\{(\ell_i \theta_i)^{\frac{1}{2}} + (\ell_i \theta_i + 4)^{\frac{1}{2}}\}$$
$$+ \tfrac{1}{2}(\ell_j \theta_j)^{\frac{1}{2}}\{(\ell_j \theta_j)^{\frac{1}{2}} + (\ell_j \theta_j + 4)^{\frac{1}{2}}\} \qquad (1 \le i < j \le k) ,$$

$$\frac{\partial^2 \psi}{\partial s_{ij} \partial w_{ij}} = \tfrac{1}{2}(\theta_j \ell_i)^{\frac{1}{2}}\{(\ell_i \theta_i)^{\frac{1}{2}} + (\ell_i \theta_i + 4)^{\frac{1}{2}}\} + \tfrac{1}{2}(\theta_i \ell_j)^{\frac{1}{2}}\{(\ell_j \theta_j)^{\frac{1}{2}} + (\ell_j \theta_j + 4)^{\frac{1}{2}}\}$$
$$(1 \le i < j \le k) ,$$

$$\frac{\partial^2 \psi}{\partial s_{ij} \partial z_{ij}} = \tfrac{1}{2}(\theta_j \ell_j)^{\frac{1}{2}}\{(\ell_i \theta_i)^{\frac{1}{2}} + (\ell_i \theta_i + 4)^{\frac{1}{2}}\} + \tfrac{1}{2}(\theta_i \ell_i)^{\frac{1}{2}}\{(\ell_j \theta_j)^{\frac{1}{2}} + (\ell_j \theta_j + 4)^{\frac{1}{2}}\}$$
$$(1 \le i < j \le k) ,$$

$$\frac{\partial^2 \psi}{\partial w_{ij} \partial z_{ij}} = \tfrac{1}{2}(\theta_i \ell_j)^{\frac{1}{2}}\{(\ell_i \theta_i)^{\frac{1}{2}} + (\ell_i \theta_i + 4)^{\frac{1}{2}}\} + \tfrac{1}{2}(\theta_j \ell_i)^{\frac{1}{2}}\{(\ell_j \theta_j)^{\frac{1}{2}}$$
$$+ (\ell_j \theta_j + 4)^{\frac{1}{2}}\} \qquad (1 \le i < j \le k) ,$$

$$\frac{\partial^2 \psi}{\partial w_{ij}^2} = \frac{\partial^2 \psi}{\partial z_{ij}^2} = \tfrac{1}{2}(\theta_i \ell_i)^{\frac{1}{2}}\{(\ell_i \theta_i)^{\frac{1}{2}} + (\ell_i \theta_i + 4)^{\frac{1}{2}}\} \qquad (1 \le i \le k, k+1 \le j \le p) ,$$

$$\frac{\partial^2 \psi}{\partial w_{ij} \partial z_{ij}} = \tfrac{1}{2}(\theta_i \ell_j)^{\frac{1}{2}}\{(\ell_i \theta_i)^{\frac{1}{2}} + (\ell_i \theta_i + 4)^{\frac{1}{2}}\} \qquad (1 \le i \le k, k+1 \le j \le p) ,$$

and

$$\frac{\partial^2\psi}{\partial z_{ij}^2} = \tfrac{1}{2}(\theta_i \ell_i)^{\frac{1}{2}}\{(\ell_i \theta_i)^{\frac{1}{2}} + (\ell_i \theta_i + 4)^{\frac{1}{2}}\} \qquad (1 \le i \le k,\ p+1 \le j \le n_1) .$$

It follows from these results and the fact that $(\theta, \ell) \in B$ implies $\epsilon < \ell_i < 1$ $(1 \le i \le k)$ and $\epsilon < \theta_j < K$ $(1 \le j \le p)$ that

$$\sup_B |w_{ij}\{\xi(\theta, \ell), \theta, \ell\}| < 2(K+4) = \zeta < \infty .$$

By Lemma 2.3.1

$$\begin{aligned}
\gamma_1^2(\theta, \ell) &= \max_{x'x=1} x'\Omega\{\xi(\theta, \ell), \theta, \ell\}x \\
&\le \max_{x'x=1} \sum_{i=1}^{k} \sum_{j=1}^{k} |x_i|\,|x_j|\,|w_{ij}\{\xi(\theta, \ell), \theta, \ell\}| \\
&\le \zeta \max_{x'x=1} \{\sum_{i=1}^{k} |x_i|\}^2 \\
&\le k^2 \zeta .
\end{aligned}$$

Thus

$$\sup_B \gamma_1^2(\theta, \ell) < \infty .$$

To prove that $\inf_B \gamma_k^2(\theta, \ell) > 0$ it is sufficient, because of (3.2.25), to prove that $\inf_B \Delta\{\xi(\theta, \ell)\} > 0$, where $\Delta\{\xi(\theta, \ell)\}$ is the Hessian of ψ as a function x evaluated at $\xi(\theta, \ell)$.

By interchanging pairs of rows and the corresponding pairs of columns of $\Omega\{\xi(\theta, \ell), \theta, \ell\}$, we can express $\Delta\{\xi(\theta, \ell)\}$ as the determinant of a block diagonal matrix of the form

$$\mathrm{diag}(W_1, \ldots, W_k, X_{12}, \ldots, X_{k-1,k}, Y_{1,k+1}, \ldots, Y_{1p}, \ldots, Y_{k,k+1}, \ldots, Y_{kp}, Z_1, \ldots, Z_k)$$

where

$$W_i = \frac{\partial^2 \psi}{\partial v_i^2} \qquad (1 \le i \le k) ,$$

$$X_{ij} = \begin{bmatrix} \frac{\partial^2 \psi}{\partial s_{ij}^2} & \frac{\partial^2 \psi}{\partial s_{ij} \partial w_{ij}} & \frac{\partial^2 \psi}{\partial s_{ij} \partial z_{ij}} \\ \frac{\partial^2 \psi}{\partial w_{ij} \partial s_{ij}} & \frac{\partial^2 \psi}{\partial w_{ij}^2} & \frac{\partial^2 \psi}{\partial w_{ij} \partial z_{ij}} \\ \frac{\partial^2 \psi}{\partial z_{ij} \partial s_{ij}} & \frac{\partial^2 \psi}{\partial z_{ij} \partial w_{ij}} & \frac{\partial^2 \psi}{\partial z_{ij}^2} \end{bmatrix} \qquad (1 \le i < j \le k) ,$$

$$Y_{ij} = \begin{bmatrix} \frac{\partial^2 \psi}{\partial w_{ij}^2} & \frac{\partial^2 \psi}{\partial w_{ij} \partial z_{ij}} \\ \frac{\partial^2 \psi}{\partial z_{ij} \partial w_{ij}} & \frac{\partial^2 \psi}{\partial z_{ij}^2} \end{bmatrix} \qquad (1 \le i \le k, k+1 \le j \le p) ,$$

and

$$Z_i = \tfrac{1}{2}(\ell_i \theta_i)^{\frac{1}{2}}\{(\ell_i \theta_i)^{\frac{1}{2}} + (\ell_i \theta_i + 4)^{\frac{1}{2}}\} I_{n_1 - p} .$$

All of the derivatives are evaluated at the point $\{\xi(\theta, \ell), \theta, \ell\}$. We then have

$$\Delta\{\xi(\theta, \ell)\} = \prod_{i=1}^{k} W_i \prod_{i<j}^{k} \det X_{ij} \prod_{i=1}^{k} \prod_{j=k+1}^{p} \det Y_{ij} \prod_{i=1}^{k} \det Z_i .$$

After a considerable amount of simplification $\Delta\{\xi(\theta, \ell)\}$ reduces to

$$(5.2.17) \quad \Delta\{\xi(\theta, \ell)\} = 2^{k(k+2-p-n_1)} \prod_{i=1}^{k} \prod_{j=k+1}^{p} [\{(\ell_i \theta_i)^{\frac{1}{2}} + (\ell_i \theta_i + 4)^{\frac{1}{2}}\}^2 (\ell_i - \ell_j) \theta_i]$$

$$\times \prod_{i<j}^{k} [(\ell_i - \ell_j)(\theta_i - \theta_j)\{\ell_i \theta_i - \ell_j \theta_j + (\ell_i \theta_i)^{\frac{1}{2}} (\ell_i \theta_i + 4)^{\frac{1}{2}} - (\ell_j \theta_j)^{\frac{1}{2}} (\ell_j \theta_j + 4)^{\frac{1}{2}}\}^2]$$

$$\times \prod_{i=1}^{k} [(\theta_i \ell_i)^{\frac{1}{2}(n_1 - p)} \{(\ell_i \theta_i)^{\frac{1}{2}} + (\ell_i \theta_i + 4)^{\frac{1}{2}}\}^{n_1 - p - 1} (\ell_i \theta_i + 4)^{\frac{1}{2}}] .$$

It follows from (5.2.17) and the fact that $(\theta, \ell) \in B$ implies $\theta_i - \theta_{i+1} > \epsilon$ and $\ell_j - \ell_{j+1} > \epsilon$ $(1 \le i \le k, 1 \le j \le p)$ where $\theta_{k+1} = \ell_{p+1} = 0$, that

$$\inf_B \Delta\{\xi(\theta, \ell)\} > 0 .$$

(vii) Ω_i was defined in (5.2.6) as

$$\Omega_i = N_{i1} \times N_{i2} \times D_V \times N_{i3}$$

where $N_{i1} \subset V(k,p)$, $N_{i2} \subset O(k)$, and $N_{i3} \subset V(k,n_1)$ are neighborhoods of

$$\begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix}, \quad I_k, \quad \text{and} \quad \begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix}, \text{ respectively.}$$

If $y \in \Omega_i$ then the components of y are values of v, G, H_1, and E_1. Write

$$y = (H_1, G, V, E_1)$$

and let

$$y_A = \left(\begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix}, \; I_k, \; v, \; \begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix} \right) .$$

We can verify condition (vii) by the same argument that was used in Chapter 3 to verify condition (vii) once we have proved the following lemma.

Lemma 5.2.2--

For every $\delta > 0$ there exists a constant $\mu = \mu(\delta)$, $0 < \mu < 1$, such that if $y \in \Omega_i$ for some i $(1 \le i \le 2^{2k})$ and $d(y_A, \beta(\theta, \ell)) \ge \delta$ for some $(\theta, \ell) \in B$, then

$$|h(y, \theta, \ell)/h\{\beta(\theta, \ell), \theta, \ell\}| < \mu .$$

Proof--

Let $t_j = t_j(\theta, \ell) = \frac{1}{2}\{(\theta_j \ell_j)^{\frac{1}{2}} + (\theta_j \ell_j + 4)^{\frac{1}{2}}\}$ $(1 \leq j \leq k)$. It follows from the definitions of y_A and $\beta(\theta, \ell)$ that

$$d(y_A, \beta(\theta, \ell)) = \left\{ \sum_{j=1}^{k} (v_j - t_j)^2 \right\}^{\frac{1}{2}} .$$

That there exists a J $(1 \leq J \leq k)$ such that $|v_J - t_J| \geq \delta k^{-\frac{1}{2}}$ is implied by $d(y_A, \beta(\theta, \ell)) \geq \delta$. From Lemma (5.2.1) and Corollary 1 of the appendix

$$\left| \frac{h(y, \theta, \ell)}{h\{\beta(\theta, \ell), \theta, \ell\}} \right| \leq \sup_{G, H_1, E_1} \left| \frac{h(y, \theta, \ell)}{h\{\beta(\theta, \ell), \theta, \ell\}} \right|$$

$$= \exp\left[\sum_{i=1}^{k} \{ -\tfrac{1}{2} v_i^2 + (\theta_i \ell_i)^{\frac{1}{2}} v_i + \ln v_i - \ln t_i + \tfrac{1}{2} - \tfrac{1}{2}(\theta_i \ell_i)^{\frac{1}{2}} t_i \} \right]$$

$$\leq \exp\{g_1(v_J, \theta_J, \ell_J)\}$$

where

$$g_1(v_J, \theta_J, \ell_J) = -\tfrac{1}{2} v_J^2 + (\theta_J \ell_J)^{\frac{1}{2}} v_J + \ln v_J - \ln t_J + \tfrac{1}{2} - \tfrac{1}{2}(\theta_J \ell_J)^{\frac{1}{2}} t_J .$$

By Cauchy's inequality

$$v_J = v_J t_J t_J^{-1} \leq \tfrac{1}{2} t_J^{-1} (v_J^2 + t_J^2) .$$

Therefore $g_1(v_J, \theta_J, \ell_J) \leq g_2(v_J, \theta_J, \ell_J)$ where

$$g_2(v_J, \theta_J, \ell_J) = -\tfrac{1}{2} v_J^2 \{1 - t_J^{-1} (\theta_J \ell_J)^{\frac{1}{2}}\} + \ln v_J - \ln t_J + \tfrac{1}{2}$$

$$= -\tfrac{1}{2} v_J^2 t_J^{-2} + \ln(v_J t_J^{-1}) + \tfrac{1}{2} .$$

To prove the lemma it suffices to prove that

$$\sup_{v_J, \theta_J, \ell_J} g_2(v_J, \theta_J, \ell_J) = \nu < 0 .$$

The lemma will then follow by taking $\theta = e^{\nu}$. Since $|v_J - t_J| \geq \delta k^{-\frac{1}{2}}$ we can write

$$v_J = t_J + d$$

where $|d| \geq \delta k^{-\frac{1}{2}}$ and $t_J + d > 0$. Make this change of variables in g_2. After simplifying we have

$$g_2(v_J, \theta_J, \ell_J) = g_3(dt_J^{-1})$$

where

$$g_3(t) = -\tfrac{1}{2}t^2 - t + \ln(1+t), \qquad t > -1 .$$

It is easily shown that g_3 has a maximum value of 0 which is obtained at $t = 0$, $g_3'(t) > 0$ $(-1<t<0)$, and $g_3'(t) < 0$ $(0<t)$. Since $t_J + d > 0$ we have $dt_J^{-1} > -1$. If $(\theta, \ell) \in B$ then $\epsilon < \theta_J < K$, $\epsilon < \ell_J < 1$, and therefore

$$t_J^{-1} > 2\{K^{\frac{1}{2}} + (K+4)^{\frac{1}{2}}\}^{-1} = X > 0 .$$

Combining these results we have

$$\sup_{v_J, \theta_J, \ell_J} g_2(v_J, \theta_J, \ell_J) = \sup_{dt_J^{-1}} g_3(dt_J^{-1})$$

$$\leq \max\{g_3(\delta k^{-\frac{1}{2}}X), g_3(-\delta k^{-\frac{1}{2}}X)\}$$

$$< g(0) = 0$$

proving the lemma.

The argument used in Chapter 3 may now be applied. Summarizing we have

<u>Lemma 5.2.3</u>--

For every $\delta > 0$ there exists a constant $\mu = \mu(\delta)$, $0 < \mu < 1$, such that

$$\left|\frac{h(y,\theta,\ell)}{h[\beta(\theta,\ell),\theta,\ell]}\right| < \mu$$

for all $y \in \Omega_i - S[\delta, \beta(\theta,\ell)]$, $(\theta,\ell) \in B$, and i $(1 \le i \le 2^{2k})$.

Since all of the conditions of Theorem 2.3.1 are satisfied we have

(5.2.18) $L_i \sim \tau[\xi(\theta,\ell)][f(\xi(\theta,\ell),\theta,\ell)]^{n_2}[\Delta(\xi(\theta,\ell))]^{-\frac{1}{2}}(2\pi/n_2)^{\frac{1}{2}k\{n_1+p-\frac{1}{2}(k+1)\}}$

uniformly in $(\theta,\ell) \in B$. Using (5.2.17), (5.2.18) reduces to

(5.2.19) $$L_i \sim C_2 \prod_{i<j}^{k} [(\theta_i-\theta_j)(\ell_i-\ell_j)]^{-\frac{1}{2}} \prod_{i=1}^{k} \prod_{j=k+1}^{p} \{\theta_i(\ell_i-\ell_j)\}^{-\frac{1}{2}}$$

$$\times \prod_{i=1}^{k} [\{(\theta_i\ell_i)^{\frac{1}{2}}+(\theta_i\ell_i+4)^{\frac{1}{2}}\}^{n_2+\frac{1}{2}(n_1-p+1)} (\ell_i\theta_i+4)^{-\frac{1}{4}}(\ell_i\theta_i)^{(p-n_1)/4}] \cdot$$

$$\times \exp\left[\frac{1}{4} n_2 \sum_{i=1}^{k} [(\ell_i\theta_i)^{\frac{1}{2}}\{(\ell_i\theta_i)^{\frac{1}{2}}+(\ell_i\theta_i+4)^{\frac{1}{2}}\}]\right]$$

where

$$C_2 = \frac{\Gamma_k(\frac{1}{2}k)\Gamma_k(\frac{1}{2}p)\Gamma_k(\frac{1}{2}n_1)e^{-\frac{1}{2}n_2k}\, 2^{\frac{1}{2}k\{2p-2n_2-\frac{1}{2}k-15/2\}}}{\pi^{k(3k+1)/4}\, n_2^{\frac{1}{2}k\{n_1+p-\frac{1}{2}k-\frac{1}{2}\}}}$$

<u>Asymptotic behavior of L_{i2}</u>--

By exactly the same argument that was used in Chapter 3, we can show that

$$|L_{i2}/\hat{L}_1| \le \mu^{n_2+n_1-k}\, n_2^{\frac{1}{2}k\{n_1+p-\frac{1}{2}(k+1)\}}\, \iota$$

for all $y \in \Omega_i - \Xi$, i $(1 \le i \le 2^{2k})$, and $(\theta,\ell) \in B$, where ι is a constant, μ is a constant, $0 < \mu < 1$, such that $|h(y,\theta,\ell)/h[\beta(\theta,\ell),\theta,\ell]| < \mu$, and $\hat{L}_1$ is given by the righthand side of (5.2.18). Therefore L_{i2} is asymptotically of lower order of magnitude than L_1 .

The results of this section are summarized in the following theorem.

Theorem 5.2.1.

Let

i) $L = \mathrm{diag}(\ell_1, \ldots, \ell_p)$ with $1 > \ell_1 > \cdots > \ell_p > 0$;

ii) $\Theta = \mathrm{diag}(\theta_1, \ldots, \theta_p)$ with $\theta_1 > \theta_2 > \cdots > \theta_k > \theta_{k+1} = \cdots = \theta_p = 0$;

iii) $A = \{(\theta, \ell) : 1 > \ell_1 > \cdots > \ell_p > 0 ,\ \theta_1 > \theta_2 > \cdots > \theta_k > 0\}$
where $\theta = (\theta_1, \ldots, \theta_k)'$ and $\ell = (\ell_1, \ldots, \ell_p)'$; and

iv) for every ϵ, $0 < \epsilon < \frac{1}{2}$, and $K > 0$,

$B = B(K, \epsilon) = \{(\theta, \ell) \in A : K > \theta_1,\ \theta_i - \theta_{i+1} > \epsilon,\ (1 \le i \le k)$ and $\ell_j - \ell_{j+1} > \epsilon,\ (1 \le j \le p)$ where $\theta_{k+1} = \ell_{p+1} = 0\}$. Then for all $(\theta, \ell) \in A$

$$(5.2.20)\quad {}_1F_1\left(\tfrac{1}{2}(n_1+n_2), \tfrac{1}{2}n_1 ; \tfrac{1}{2}n_2\Theta, L\right) \sim K_n \prod_{i=1}^{k} \prod_{j=1}^{p} \{(\theta_i - \theta_j)(\ell_i - \ell_j)\}^{-\frac{1}{2}}$$

$$\times \prod_{i=1}^{k} \left[\{(\theta_i \ell_i)^{\frac{1}{2}} + (\theta_i \ell_i + 4)^{\frac{1}{2}}\}^{n_2 + \frac{1}{2}(n_1 - p + 1)} (\ell_i \theta_i + 4)^{-\frac{1}{4}} (\ell_i \theta_i)^{(p-n_1)/4}\right]$$

$$\times \exp\left[\frac{1}{4} n_2 \sum_{i=1}^{k} [(\ell_i \theta_i)^{\frac{1}{2}} \{(\ell_i \theta_i)^{\frac{1}{2}} + (\ell_i \theta_i + 4)^{\frac{1}{2}}\}]\right]$$

where

$$K_n = \frac{\Gamma_k(\frac{1}{2}n_1)\Gamma_k(\frac{1}{2}p) n_2^{\frac{1}{2}k(n_2 - p + \frac{1}{2} + \frac{1}{2}k)}\, 2^{\frac{1}{2}k(2p - \frac{1}{2}k - 3n_2 - n_1 - 3/2)}\, e^{-\frac{1}{2}n_2 k}}{\Gamma_k\{\frac{1}{2}(n_1+n_2)\}\, \pi^{k(k+1)/4}} .$$

Furthermore (5.2.20) holds uniformly in $(\theta, \ell) \in B$ for every ϵ and K .

5.3. The Term of Order n_2^{-1} in the Asymptotic Expansion of ${}_1F_1\left(\frac{1}{2}(n_1+n_2); \frac{1}{2}n_1 ; \frac{1}{2}n_2\Theta, L\right)$.

As indicated in Chapter 2, Section 4, further terms in an asymptotic expansion of ${}_1F_1(n_2, \Theta, L)$ could be obtained, at least in principle, by a more careful analysis of the integrals in Section 5.2. Such an analysis would lead to an asymptotic expansion of the form

$$ {}_1F_1(n_2,\Theta,L) \sim \varphi G $$

where Ψ is the asymptotic representation for ${}_1F_1(n_2,\Theta,L)$ given by Theorem 5 2.1,

$$ G = 1 + \frac{P_1}{n_2} + \frac{P_2}{n_2^{\ 2}} + \cdots $$

and the P_i's are functions of the θ_i's and ℓ_i's but do not depend on n_2.

Suppose that S is a $p \times p$ symmetric matrix with latent roots $s_1,\ldots,s_p$, $T = \mathrm{diag}(T_1,0)$ is a $p \times p$ matrix, and T_1 is a $k \times k$ symmetric matrix with latent roots $t_1,\ldots,t_k$. Then by Theorem 3.3.2, ${}_1F_1(a;c;T,S)$ satisfies the following partial differential equation

$$ \sum_{i=1}^{p} s_i \frac{\partial^2 F}{\partial s_i^{\ 2}} + \sum_{i=1}^{p} \sum_{\substack{j=1 \\ i\neq j}}^{p} \frac{s_i}{s_i - s_j} \frac{\partial F}{\partial s_i} + \{c - \tfrac{1}{2}(p-1)\} \sum_{i=1}^{p} \frac{\partial F}{\partial s_i} $$

$$ - \sum_{i=1}^{k} t_i^{\ 2} \frac{\partial F}{\partial t_i} = aF \sum_{i=1}^{k} t_i . $$

Let $S = L$, $T = \frac{1}{2}n_2\Theta$, $a = \frac{1}{2}(n_1+n_2)$, and $c = \frac{1}{2}n_1$. It follows after some simplification that ${}_1F_1(\Theta,L)$ satisfies

$$ \sum_{i=1}^{p} \ell_i \frac{\partial^2 F}{\partial \ell_i^2} + \sum_{i=1}^{p} \sum_{\substack{j=1 \\ i\neq j}}^{p} \frac{\ell_i}{\ell_i - \ell_j} \frac{\partial F}{\partial \ell_i} + \tfrac{1}{2}(n_1-p+1) \sum_{i=1}^{p} \frac{\partial F}{\partial \ell_i} \tag{5.3.1} $$

$$ - \tfrac{1}{2}n_2 \sum_{i=1}^{k} \theta_i^{\ 2} \frac{\partial F}{\partial \theta_i} = \frac{n_2(n_1+n_2)}{4} F \sum_{i=1}^{k} \theta_i . $$

Substituting φG for F in (5.3.1) we have after a considerable amount of computation that G satisfies

$$\sum_{i=1}^{p} \ell_i \frac{\partial^2 G}{\partial \ell_i^2} + \sum_{i=k+1}^{p} \sum_{\substack{j=k+1 \\ i \neq j}}^{p} \frac{\ell_i}{\ell_i - \ell_j} \frac{\partial G}{\partial \ell_i} + \tfrac{1}{2}(n_1 - p + 1) \sum_{i=k+1}^{p} \frac{\partial G}{\partial \ell_i}$$

$$+ \tfrac{1}{2} n_2 \sum_{i=1}^{k} (\ell_i \theta_i)^{\frac{1}{2}} \{(\ell_i \theta_i)^{\frac{1}{2}} + (\ell_i \theta_i + 4)^{\frac{1}{2}}\} \frac{\partial G}{\partial \ell_i}$$

$$+ \tfrac{1}{2} \sum_{i=1}^{k} \left[\frac{(n_1 - p)(\ell_i \theta_i)^{\frac{1}{2}}}{(\ell_i \theta_i + 4)^{\frac{1}{2}}} + \frac{(\ell_i \theta_i)^{\frac{1}{2}} (\ell_i \theta_i + 4)^{\frac{1}{2}} + 4}{\ell_i \theta_i + 4} \right] \frac{\partial G}{\partial \ell_i}$$

$$- \tfrac{1}{2} n_2 \sum_{i=1}^{k} \theta_i^2 \frac{\partial G}{\partial \theta_i} = \frac{G}{4} \left[\sum_{i=1}^{k} \left\{ \frac{(p - n_1)(p - n_1 + 2)}{\ell_i (\ell_i \theta_i + 4)} \right. \right.$$

$$+ \frac{(p - n_1)\theta_i \{(\ell_i \theta_i + 4)^{\frac{1}{2}} - (\ell_i \theta_i)^{\frac{1}{2}}\}}{(\ell_i \theta_i + 4)^{3/2}} + \left. \frac{\theta_i \{(\ell_i \theta_i)^{\frac{1}{2}} (\ell_i \theta_i + 4)^{\frac{1}{2}} + 1 - \ell_i \theta_i\}}{(\ell_i \theta_i + 4)^2} \right\}$$

$$\left. - \sum_{i \neq j}^{k} \sum \frac{\ell_i}{(\ell_i - \ell_j)^2} - \sum_{i=1}^{k} \sum_{j=k+1}^{p} \frac{\ell_i + \ell_j}{(\ell_i - \ell_j)^2} \right] .$$

The differential equation for P_1 is obtained by substituting

$$1 + \frac{P_1}{n_2} + \frac{P_2}{n_2^2} + \cdots$$

for G in the previous differential equation and equating coefficients of powers of n_2^0. The resulting equation for P_1 is

$$(5.3.2) \quad \sum_{i=1}^{k} (\ell_i \theta_i)^{\frac{1}{2}} \{(\ell_i \theta_i + 4)^{\frac{1}{2}} + (\ell_i \theta_i)^{\frac{1}{2}}\} \frac{\partial P_1}{\partial \ell_i} - \sum_{i=1}^{k} \theta_i^2 \frac{\partial P_1}{\partial \theta_i}$$

$$= \tfrac{1}{2} P_1 \left[\sum_{i=1}^{k} \left\{ \frac{(p - n_1)(p - n_1 + 2)}{\ell_i (\ell_i \theta_i + 4)} + \frac{(p - n_1)\theta_i \{(\ell_i \theta_i + 4)^{\frac{1}{2}} - (\ell_i \theta_i)^{\frac{1}{2}}\}}{(\ell_i \theta_i + 4)^{3/2}} \right. \right.$$

$$+ \left. \frac{\theta_i \{(\ell_i \theta_i)^{\frac{1}{2}} (\ell_i \theta_i + 4)^{\frac{1}{2}} + 1 - \ell_i \theta_i\}}{(\ell_i \theta_i + 4)^2} \right\} - \sum_{i \neq j}^{k} \sum \frac{\ell_i}{(\ell_i - \ell_j)^2}$$

$$- \sum_{i=1}^{k} \sum_{j=k+1}^{p} \frac{\ell_i + \ell_j}{(\ell_i - \ell_j)^2} \Bigg] .$$

Also P_1 satisfies the boundary condition

$$P_1(\Theta, L) = P_1(L, \Theta) \tag{5.3.3}$$

since ${}_1F_1(n_2, \Theta, L) = {}_1F_1(n_2, L, \Theta)$ and $\varphi(n_2, \Theta, L) = \varphi(n_2, L, \Theta)$.

The general solution of (5.3.2) is

$$P_1 = Q + \Psi(u_1, \ldots, u_{p+k-1})$$

where Q is any particular solution, Ψ is an arbitrary function and the u_i's are any $p + k - 1$ independent solutions of the system of equations

$$\frac{d\ell_1}{\ell_1\theta_1 + (\ell_1\theta_1)^{\frac{1}{2}}(\ell_1\theta_1+4)^{\frac{1}{2}}} = \cdots = \frac{d\ell_k}{\ell_k\theta_k + (\ell_k\theta_k)^{\frac{1}{2}}(\ell_k\theta_k+4)^{\frac{1}{2}}}$$

$$= - \frac{d\theta_1}{\theta_1^{\,2}} = \cdots = - \frac{d\theta_k}{\theta_k^{\,2}} .$$

The u_i's are easily found to be

$$u_i = \theta_i^{-1} - \theta_k^{-1} \qquad i=1, \ldots, k-1$$

$$u_i = \ell_{i+1} \qquad i=k, \ldots, p-1 \tag{5.3.4}$$

$$u_i = \frac{\theta_i^{\frac{1}{2}}}{(\ell_i\theta_i)^{\frac{1}{2}} + (\ell_i\theta_i+4)^{\frac{1}{2}}} \qquad i=p, \ldots, p+k-1$$

Ψ is evaluated using the boundary conditions. If we can find a Q which satisfies the boundary conditions then so must Ψ . Examination of (5.3.4) shows that a function satisfying (5.3.3) must be identically constant. Since there is only one boundary condition for P_1 , there is no way to determine the constant.

A particular solution of (5.3.2) is

$$(5.3.5)\quad Q = \sum_{i=1}^{k}\left[-\frac{(p-n_1)^2(\ell_i\theta_i+2)}{8\,(\ell_i\theta_i)^{\frac{1}{2}}(\ell_i\theta_i+4)^{\frac{1}{2}}} + \frac{(p-n_1)\{(\ell_i\theta_i)^{\frac{1}{2}} - (\ell_i\theta_i+4)^{\frac{1}{2}}\}}{2(\ell_i\theta_i)^{\frac{1}{2}}(\ell_i\theta_i+4)} + \frac{6\,(\ell_i\theta_i)^{\frac{1}{2}} - (\ell_i\theta_i)^{3/2} - 12(\ell_i\theta_i+4)^{\frac{1}{2}}}{(\ell_i\theta_i+4)^{3/2}}\right]$$

$$+ \sum_{i<j}^{k}\sum \frac{(\ell_i\theta_i)^{\frac{1}{2}}(\ell_j\theta_j+4)^{\frac{1}{2}} + (\ell_j\theta_j)^{\frac{1}{2}}(\ell_i\theta_i+4)^{\frac{1}{2}}}{(\ell_i-\ell_j)(\theta_i-\theta_j)\{(\ell_i\theta_i)^{\frac{1}{2}} + (\ell_i\theta_i+4)^{\frac{1}{2}}\}\{(\ell_j\theta_j)^{\frac{1}{2}} + (\ell_j\theta_j+4)^{\frac{1}{2}}\}}$$

$$+ \sum_{i=1}^{k}\sum_{j=k+1}^{p} \frac{(\ell_i\theta_i)^{\frac{1}{2}}}{\theta_i(\ell_i-\ell_j)\{(\ell_i\theta_i)^{\frac{1}{2}} + (\ell_i\theta_i+4)^{\frac{1}{2}}\}}\,.$$

Note that $Q(\Theta, L) = Q(L, \Theta)$. Hence the general solution of (5.3.2) is

$$P_1 = Q + \eta$$

where Q is given by (5.3.5) and η is a constant.

Additional P_i's could be calculated by the same technique but the partial differential equations for P_i are very complicated when $i \geq 2$.

CHAPTER 6

MANOVA AND DISCRIMINANT ANALYSIS

6.1 Introduction.

Suppose W and B are independent $p \times p$ random matrices such that W has a Wishart distribution on n_2 $(n_2 \geq p)$ degrees of freedom, $W_p(n_2|\Sigma)$, and B has a noncentral Wishart distribution on n_1 $(n_1 \geq p)$ degrees of freedom and noncentrality matrix $\Sigma^{-1}\Lambda$, $W_p(n_1|\Sigma|\Sigma^{-1}\Lambda)$. Let $\Omega = \mathrm{diag}(\omega_1, \ldots, \omega_p)$, $\omega_1 \geq \omega_2 \geq \cdots \geq \omega_p \geq 0$, be the diagonal matrix of the latent roots of $\Sigma^{-1}\Lambda$ and let $L = \mathrm{diag}(\ell_1, \ldots, \ell_p)$, $\ell_1 > \ell_2 > \cdots > \ell_p > 0$, be the diagonal matrix of the latent roots of $B(B+W)^{-1}$. In MANOVA and discriminant analysis B and W are respectively the "between groups" and "within groups" matrices of sums of squares and sums of products (see Chapter 1, Section 4). Finally assume that $\Omega = n_2\Theta$, where $\Theta = \mathrm{diag}(\theta_1, \ldots, \theta_p)$ with $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_p \geq 0$. The joint pdf of $\ell_1, \ldots, \ell_p$ is (see (1.4.1))

$$(6.1.1) \quad f(L) = \mathrm{etr}(-\tfrac{1}{2}n_2\Theta)\,{}_1F_1\{\tfrac{1}{2}(n_1+n_2); \tfrac{1}{2}n_1; \tfrac{1}{2}n_2\Theta, L\}$$

$$\times C_1 \prod_{i=1}^{p} \{\ell_i^{\frac{1}{2}(n_1-p-1)} (1-\ell_i)^{\frac{1}{2}(n_2-p-1)}\} \prod_{i<j}^{p} (\ell_i - \ell_j) \prod_{i=1}^{p} d\ell_i$$

where

$$C_1 = \frac{\pi^{\frac{1}{2}p^2}\Gamma_p\{\tfrac{1}{2}(n_1+n_2)\}}{\Gamma_p(\tfrac{1}{2}n_1)\Gamma_p(\tfrac{1}{2}n_2)\Gamma_p(\tfrac{1}{2}p)}.$$

Since the distribution of $\ell_1, \ldots, \ell_p$ depends only on $\theta_1, \ldots, \theta_p$ that part of the pdf which involves $\theta_1, \ldots, \theta_p$ can be regarded as a marginal likelihood L. Then

(6.1.2) $$L = \operatorname{etr}(-\tfrac{1}{2}n_2\Theta)\,{}_1F_1\{\tfrac{1}{2}(n_1+n_2);\tfrac{1}{2}n_1;\tfrac{1}{2}n_2\Theta, L\} .$$

The results of Chapter 5 are used to derive an asymptotic expansion for large n_2 of the pdf (6.1.1) and an asymptotic representation for large n_2 of the likelihood function (6.1.2) under the assumption that the θ_i's satisfy

(6.1.3) $$\theta_1 > \theta_2 > \cdots > \theta_k > \theta_{k+1} = \cdots = \theta_p = 0 .$$

These results are then used to study Bartlett's test that the residual $p - k$ θ_i's are zero.

6.2. Asymptotic Expansions.

An asymptotic expansion of the joint pdf of $\ell_1,\ldots,\ell_p$ is obtained by substituting the asymptotic expansion of ${}_1F_1\{\tfrac{1}{2}(n_1+n_2);\tfrac{1}{2}n_2;\tfrac{1}{2}n_2\Theta, L\}$ given by Theorem 5.2.1 and (5.3.5) into (6.1.1). The result is summarized in the following theorem.

Theorem 6.2.1.

An asymptotic expansion for large n_2 of the joint pdf of $\ell_1,\ldots,\ell_p$ when $\Omega = n_2\Theta$ and the ℓ_i's satisfy (6.1.3), is

(6.2.1) $$\Psi\left(1 + \frac{P_1}{n} + O(n_2^{-2})\right)$$

where

(6.2.2) $$\varphi = C_2 \prod_{i=1}^{k} \left[\{(\ell_i\theta_i)^{\frac{1}{2}} + (\ell_i\theta_i+4)^{\frac{1}{2}}\}^{n_2+\frac{1}{2}(n_1-p+1)} (\ell_i\theta_i)^{(p-n_1)/4} (\ell_i\theta_i+4)^{-\frac{1}{4}}\right]$$

$$\times \prod_{i=1}^{k} \{\ell_i^{\frac{1}{2}(n_1-p-1)} (1-\ell_i)^{\frac{1}{2}(n_2-p-1)}\} \prod_{i<j}^{k} \left(\frac{\ell_i - \ell_j}{\theta_i - \theta_j}\right)^{\frac{1}{2}}$$

$$\times \exp\left[\tfrac{1}{4}n_2 \sum_{i=1}^{k} (\theta_i\ell_i)^{\frac{1}{2}}\{(\theta_i\ell_i)^{\frac{1}{2}} + (\theta_i\ell_i+4)^{\frac{1}{2}}\}\right]$$

$$\times \prod_{i=1}^{k} \prod_{j=k+1}^{p} (\ell_i - \ell_j)^{\frac{1}{2}}$$

$$\times \prod_{i=k+1}^{p} \{\ell_i^{\frac{1}{2}(n_1-p-1)} (1-\ell_i)^{\frac{1}{2}(n_2-p-1)}\} \prod_{\substack{k+1 \\ i<j}}^{p} (\ell_i - \ell_j) ,$$

$$C_2 = \frac{\Gamma_{p-k}\{\frac{1}{2}(n_1+n_2-k)\} n_2^{\frac{1}{2}k(n_2-p+\frac{1}{2}k+\frac{1}{2})} \, 2^{\frac{1}{2}k(2p-\frac{1}{2}k-3/2-3n_2-n_1)}}{\Gamma_{p-k}\{\frac{1}{2}(n_1-k)\}\Gamma_{p-k}\{\frac{1}{2}(p-k)\}\Gamma_p(\frac{1}{2}n_2) \, \pi^{\{-p^2+k-(p-k)^2\}/4}}$$

$$\times \exp\{-\tfrac{1}{2}n_2(k + \sum_{i=1}^{k} \theta_i)\} \prod_{i=1}^{k} \theta_i^{-\frac{1}{2}(p-k)} ,$$

$$P_1 = \sum_{i=1}^{k} \left[- \frac{(p-n_1)^2(\ell_i\theta_i+2)}{8\,(\ell_i\theta_i)^{\frac{1}{2}}(\ell_i\theta_i+4)^{\frac{1}{2}}} + \frac{(p-n_1)\{(\ell_i\theta_i)^{\frac{1}{2}} - (\ell_i\theta_i+4)^{\frac{1}{2}}\}}{2(\ell_i\theta_i)^{\frac{1}{2}}(\ell_i\theta_i+4)} \right.$$

$$\left. + \frac{6\,(\ell_i\theta_i)^{\frac{1}{2}} - (\ell_i\theta_i)^{3/2} - 12\,(\ell_i\theta_i+4)^{\frac{1}{2}}}{(\ell_i\theta_i+4)^{3/2}} \right]$$

$$+ \sum_{i<j}^{k}\sum \frac{(\ell_i\theta_i)^{\frac{1}{2}}(\ell_j\theta_j+4)^{\frac{1}{2}} + (\ell_j\theta_j)^{\frac{1}{2}}(\ell_i\theta_i+4)^{\frac{1}{2}}}{(\ell_i-\ell_j)(\theta_i-\theta_j)\{(\ell_i\theta_i)^{\frac{1}{2}} + (\ell_i\theta_i+4)^{\frac{1}{2}}\}\{(\ell_j\theta_j)^{\frac{1}{2}} + (\ell_j\theta_j+4)^{\frac{1}{2}}\}}$$

$$+ \sum_{i=1}^{k} \sum_{j=k+1}^{p} \frac{(\ell_i\theta_i)^{\frac{1}{2}}}{\theta_i(\ell_i-\ell_j)\{(\ell_i\theta_i)^{\frac{1}{2}} + (\ell_i\theta_i+4)^{\frac{1}{2}}\}} + \eta ,$$

η is a constant, and $O(n^{-2})$ are terms which for fixed ℓ_i's and θ_i's are $O(n^{-2})$ as $n \to \infty$

Furthermore, for every $\epsilon > 0$ and $K > 0$ let $B(K,\epsilon)$ be the set

(6.2.3) $B(K,\epsilon) = \{(\ell_1,\ldots,\ell_p,\theta_1,\ldots,\theta_k) : K > \theta_1 , \quad \theta_i - \theta_{i+1} > \epsilon \quad (1 \leq i \leq k)$

and $\quad \ell_j - \ell_{j+1} > \epsilon \quad (1 \leq j \leq p) \quad$ where $\quad \theta_{k+1} = \ell_{p+1} = 0\}$.

Then $f(L) \sim \varphi$ uniformly on $B(K,\epsilon)$ for every $K > 0$ and $\epsilon > 0$.

The function φ is called the "asymptotic pdf" or "asymptotic distribution" of $\ell_1,\ldots,\ell_p$.

The following corollaries follow easily from Theorem 6.2.1.

Corollary 6.2.1--

An asymptotic representation for large n_2 of the marginal likelihood function L defined by (6.1.2), when the ℓ_i's satisfy (6.1.3), is

$$\hat{L} = C_3 \prod_{i=1}^{k} [\{(\ell_i\theta_i)^{\frac{1}{2}} + (\ell_i\theta_i+4)^{\frac{1}{2}}\}^{n_2+\frac{1}{2}(n_1-p+1)} \theta_i^{(2k-p-n_1)/4}(\ell_i\theta_i+4)^{-\frac{1}{4}}]$$

$$\times \exp[\tfrac{1}{2}n_2 \sum_{i=1}^{k} \{\tfrac{1}{2}(\ell_i\theta_i)^{\frac{1}{2}}\{(\ell_i\theta_i)^{\frac{1}{2}} + (\ell_i\theta_i+4)^{\frac{1}{2}}\} - \theta_i\}]$$

$$\times \prod_{i<j}^{\ell} (\theta_i-\theta_j)^{-\frac{1}{2}},$$

where C_3 is a constant which does not depend on $\theta_1,\ldots,\theta_k$. Furthermore, $L \sim \hat{L}$ uniformly on $B(K,\epsilon)$ defined by (6.2.3), for every $K > 0$ and $\epsilon > 0$.

Corollary 6.2.2.

$\ell_1,\ldots,\ell_k$ are asymptotically sufficient for $\theta_1,\ldots,\theta_k$.

Corollary 6.2.3.

The asymptotic conditional pdf for large n_2 of $\ell_{k+1},\ldots,\ell_p$ given $\ell_1,\ldots,\ell_k$ is

$$(6.2.4)\quad \varphi_c = C_4 \left\{ \prod_{i=1}^{k} \prod_{j=k+1}^{p} (\ell_i-\ell_j)^{\frac{1}{2}} \right\} \prod_{i=k+1}^{p} \{\ell_i^{\frac{1}{2}(n_1-p-1)} (1-\ell_i)^{\frac{1}{2}(n_2-p-1)}\}$$

$$\times \prod_{\substack{k+1 \\ i<j}}^{p} (\ell_i-\ell_j).$$

Furthermore if $f(\ell_{k+1},\ldots,\ell_p|\ell_1,\ldots,\ell_k)$ is the true conditional distribu-

tion of $\ell_{k+1},\ldots,\ell_p$ given $\ell_1,\ldots,\ell_k$, then $f(\ell_{k+1},\ldots,\ell_p|\ell_1,\ldots,\ell_k) \sim \varphi_c$ uniformly on $B(K,\epsilon)$ for every $K > 0$ and $\epsilon > 0$.

The asymptotic pdf φ, defined by (6.2.2), contains linkage factors of the form $\ell_i - \ell_j$. By making the transformation of variables suggested by Hsu (1941a) it is possible to obtain a "normal type" approximation which no longer contains linkage factors corresponding to distinct θ_i's. Hsu let

$$(6.2.5) \qquad z_i = n_2^{\frac{1}{2}}\sigma_i^{-1}\{\ell_i(1-\ell_i)^{-1}-\theta_i\} \qquad (1\leq i\leq k)$$

where
$$\sigma_i = (2\theta_i)^{\frac{1}{2}}(\theta_i+2)^{\frac{1}{2}},$$

and

$$(6.2.6) \qquad z_j = n_2\ell_j(1-\ell_j)^{-1} \qquad (k+1\leq j\leq p).$$

Making this change of variables in (6.2.2) and simplifying gives

<u>Corollary 6.2.4</u>.

The asymptotic joint pdf of $z_1,\ldots,z_p$, when the θ_i's satisfy (6.1.3), is

$$(6.2.7) \left\{\prod_{i=1}^{k} g(z_i)\right\} C_5 \prod_{i=k+1}^{p}\{z_i^{\frac{1}{2}(n_1-p-1)} e^{-\frac{1}{2}z_i}\} \prod_{\substack{k+1\\ i<j}}^{p}(z_i-z_j)$$

$$\times \left\{1 + n_2^{-\frac{1}{2}}\sum_{i=1}^{k}(\alpha_i z_i+\beta_i z_i^3) + O(n_2^{-1})\right\},$$

where $g(z_i)$ denotes the standard normal density,

$$C_5 = \frac{\pi^{\frac{1}{2}(p-k)^2}}{2^{\frac{1}{2}(n_1-k)(p-k)}\,\Gamma_{p-k}\{\frac{1}{2}(p-k)\}\Gamma_{p-k}\{\frac{1}{2}(n_1-k)\}},$$

$$\beta_i = (3\sigma_i)^{-1}\{4\theta_i + 4 - 2(\theta_i+2)^{-1}\},$$

and

$$\alpha_i = \sigma_i^{-1}\{n_1 - 2\theta_i - 3 - p + 2(\theta_i+2)^{-1} + (p-k)(2+\theta_i)$$

$$+ \theta_i(2+\theta_i)\sum_{\substack{j=1\\ j\neq i}}^{k}(\theta_i-\theta_j)^{-1}\} .$$

Here $O(n_2^{-1})$ means terms which are $O(n_2^{-1})$ as $n_2 \to \infty$ uniformly on any bounded set of z_i's.

The first line of (6.2.7) shows that asymptotically the z_i's which correspond to distinct nonzero θ_i's are standard normal, independent of z_j $(i \neq j)$. The z_i's corresponding to θ_i's equal to zero are dependent and their asymptotic distribution is the same as the distribution of the latent roots of a $(p-k) \times (p-k)$ matrix having the Wishart distribution $W_{p-k}(n_1-k, I_{p-k})$.

Hsu (1941a) found the limiting distribution of the z_i's when the θ_i's have arbitrary multiplicity. His result reduces to the first line of (6.2.7) when the nonzero θ_i's are distinct. Fujikoshi (1976) found the limiting distribution of the z_i's, up to and including terms of order $n_2^{-\frac{1}{2}}$, when the θ_i's have arbitrary multiplicity and are nonzero. If $p = k$ then (6.2.7) reduces to Fujikoshi's result.

6.3. Tests of Significance of the Residual θ_i .

In this section we examine Bartlett's test of the null hypothesis H_k that $\theta_{k+1} = \cdots = \theta_p = 0$ given that $\theta_1 > \theta_2 > \cdots > \theta_k > 0$. The approach followed is similar to the one used by James (1969).

Bartlett suggested that H_k could be tested with the statistic

$$L_k = \{n_2 + \tfrac{1}{2}(n_1-p-1)\}T_k$$

where

$$T_k = - \prod_{i=k+1}^{p} \ln(1-\ell_i) .$$

It is well known that $n_2 T_k$ is asymptotically distributed as χ^2 on $(n_1-k)(p-k)$ degrees of freedom. The factor $n_2 + \frac{1}{2}(n_1-p-1)$ was chosen to improve the agreement between the moments of the test statistic and the corresponding moments of $\chi^2_{(p-k)(n_1-k)}$

The θ_i's are nuisance parameters in this test. Since $\ell_1, .. , \ell_k$ are asymptotically sufficient for $\theta_1, \ldots, \theta_k$ (Corollary 6.2.2), the effect of the θ_i's can be eliminated, at least asymptotically, by using the asymptotic conditional distribution of $\ell_{k+1}, \ldots, \ell_p$ given $\ell_1, \ldots, \ell_k$. The technique used here is to find the appropriate factor by computing the mean and variance of T_k with respect to the conditional density φ_c given by (6.2.4).

There are two points worth noting about φ_c.

(i) If we make the following identifications

$$\ell_i \to r_i^2 , \quad n_1 \to q , \quad p \to p , \quad \text{and} \quad n_1 + n_2 \to n$$

in φ_c defined by (6.2.4), then we get the asymptotic conditional pdf φ_c (4.2.4) of $r_{k+1}^2, \ldots, r_p^2$ given $r_1^2, \ldots, r_k^2$. It follows from this fact that the same argument that was used in the canonical correlation case may be used to compute the mean and variance of T_k with respect to φ_c.

(ii) Constantine and Muirhead (1976) found an asymptotic representation for the joint pdf of $\ell_1, \ldots, \ell_p$ for large $\omega_1, \ldots, \omega_k$ when $\Omega = \operatorname{diag}(\omega_1, \ldots, \omega_p)$ and $\omega_1 > \omega_2 > \cdots > \omega_k > \omega_{k+1} \geq \cdots \geq \omega_p \geq 0$. While their asymptotic representation is markedly different from that given in Theorem 6.2.1, the asymptotic conditional pdf of $\ell_{k+1}, \ldots, \ell_p$

given $\ell_1, \ldots, \ell_k$ is still φ_c. In other words, φ_c serves as the asymptotic conditional distribution both when $\omega_1, \ldots, \omega_k$ are large independently of n_2 and when $\omega_1, \ldots, \omega_k$ are large in the sense that $\omega_i = n_2 \theta_i$ and n_2 is large.

The following theorem is a direct consequence of observation (i) and Theorem 4.3.1.

Theorem 6.3.1.

The statistic

$$L_k = -\{n_2 - k + \tfrac{1}{2}(n_1 - p - 1) + \sum_{i=1}^{k} \ell_i^{-1}\} \ln \prod_{i=k+1}^{p} (1 - \ell_i)$$

is approximately distributed as χ^2 on $(p-k)(n_1-k)$ degrees of freedom

If the observed values of $\ell_1, \ldots, \ell_k$ are all near one, then the multiplying factor in L_k is approximately $-\{n_2 + \tfrac{1}{2}(n_1 - p - 1)\}$, which is the value suggested by Bartlett.

APPENDIX

The primary aim of this Appendix is to prove two results (Theorem 2 and Corollary 1) about the maximization of certain matrix functions. The author's first approach to proving these results was to determine the critical points of the functions by setting the appropriate differentials equal to zero. The objection to this line of argument is that it is complicated and provides little insight. An alternative approach is based on Theorem 1. The advantage of this method of proof is that it is fairly simple and easily generalized.

<u>Theorem 1</u>--

Suppose that

(i) $U = \mathrm{diag}(u_1, \ldots, u_k)$ with $u_1 \geq u_2 \geq \cdots \geq u_k \geq 0$;

(ii) $B(\alpha) = (b_{ij}(\alpha))$ is a $k \times k$ matrix defined for all $\alpha \in A$; and

(iii) there exists constants $\delta_1 \geq \delta_2 \geq \cdots \geq \delta_k$ such that for all $\alpha \in A$

$$b_{11}(\alpha) \leq \delta_1$$

$$b_{11}(\alpha) + b_{22}(\alpha) \leq \delta_1 + \delta_2$$

$$\vdots$$

$$b_{11}(\alpha) + \cdots + b_{kk}(\alpha) \leq \delta_1 + \cdots + \delta_k .$$

Then

(a) $$\mathrm{tr}\{UB(\alpha)\} \leq \sum_{i=1}^{k} u_i \delta_i \qquad \text{for all } \alpha \in A .$$

Furthermore, if

(iv) the u_i's are distinct and nonzero, then

(b) $$\operatorname{tr}\{UB(\alpha)\} = \sum_{i=1}^{k} u_i \delta_i \quad \text{if and only if}$$

$$b_{11}(\alpha) = \delta_1$$

$$b_{11}(\alpha) + b_{22}(\alpha) = \delta_1 + \delta_2$$

$$\vdots$$

$$b_{11}(\alpha) + \cdots + b_{kk}(\alpha) = \delta_1 + \cdots + \delta_k$$

Proof--

Let

$$S_i(\alpha) = b_{11}(\alpha) + \cdot\cdot\cdot + b_{ii}(\alpha)$$

and $$\Delta_i = \delta_1 + \cdots + \delta_i \qquad (1 \le i \le k) .$$

Then

$$\operatorname{tr}\{UB(\alpha)\} = \sum_{i=1}^{k} u_i b_{ii}(\alpha)$$

$$= b_{11}(\alpha)(u_1 - u_2) + \{b_{11}(\alpha) + b_{22}(\alpha)\}(u_2 - u_3)$$

$$+ \cdots + \{b_{11}(\alpha) + \cdots + b_{k-1,k-1}(\alpha)\}(u_{k-1} - u_k)$$

$$+ \{b_{11}(\alpha) + \cdots + b_{kk}(\alpha)\}u_k$$

$$= \sum_{i=1}^{k-1} S_i(\alpha)(u_i - u_{i+1}) + S_k(\alpha)u_k .$$

Similarly

$$\sum_{i=1}^{k} u_i \delta_i = \sum_{i=1}^{k-1} \Delta_i(u_i - u_{i+1}) + \Delta_k u_k .$$

Condition (i) says that $u_k \ge 0$ and $u_i - u_{i+1} \ge 0$ $(1 \le i \le k-1)$. Condition (iii) says that $S_i(\alpha) \le \Delta_i$ for all $\alpha \in A$ and i $(1 \le i \le k)$.

Therefore

(1) $$\sum_{i=1}^{k} u_i \delta_i - \operatorname{tr}\{UB(\alpha)\} = \sum_{i=1}^{k-1} \{\Delta_i - S_i(\alpha)\}(u_i - u_{i+1})$$

$$+ \{\Delta_k - S_k(\alpha)\}u_k \ge 0 ,$$

proving (a). If the u_i's are distinct and nonzero, then equality holds in (1) if and only if $\Delta_i = S_i(\alpha)$ $(1 \le i \le k)$, proving (b).

Lemma 1--

Assume that

(i) $H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \in O(m)$ where H_{11} is $r \times s$;

(ii) α is an $s \times 1$ vector such that $\alpha'\alpha \le C$; and

(iii) $\beta = H_{11}\alpha$.

Then

$$\beta'\beta \le C .$$

Proof--

Let γ and η be $m \times 1$ vectors defined by $\gamma' = (\alpha', 0')$, where 0 is a $(m-s) \times 1$ vector of zeros, and $\eta = H\gamma$. Then

$$\eta'\eta = (H\gamma)'(H\gamma) = \gamma'H'H\gamma = \gamma'\gamma = \alpha'\alpha \le C .$$

Also $\eta = \begin{bmatrix} H_{11}\alpha \\ H_{21}\alpha \end{bmatrix}$ and therefore

$$\begin{aligned} \eta'\eta &= \alpha'H_{11}'H_{11}\alpha + \alpha'H_{21}'H_{21}\alpha \\ &= \beta'\beta + (H_{21}\alpha)'(H_{21}\alpha) \\ &\ge \beta'\beta . \end{aligned}$$

Combining these results

$$\beta'\beta \le C .$$

Theorem 2--

Suppose that

(i) $U = \mathrm{diag}(u_1, \ldots, u_k)$ with $u_1 \ge u_2 \ge \cdots \ge u_k \ge 0$;

(ii) $V = \mathrm{diag}(v_1, \ldots, v_k)$ with $v_1 \ge v_2 \ge \cdots \ge v_k \ge 0$;

(iii) $P_1 = \mathrm{diag}(\rho_1,\ldots,\rho_k)$ with $\rho_1 \geq \rho_2 \geq \cdots \geq \rho_k \geq 0$;

(iv) $R = \mathrm{diag}(r_1,\ldots,r_p)$ with $r_1 \geq r_2 \geq \cdots \geq r_p \geq 0$, $(p \geq k)$;

(v) $F = (f_{ij}) \in O(k)$ and $G = (g_{ij}) \in O(k)$ $(1 \leq i, j \leq k)$;

(vi) $H_1 = (h_{ij}) \in V(k,p)$ $(1 \leq i \leq p, 1 \leq j \leq k)$;

(vii) $E_1 = (e_{ij}) \in V(k,q)$ $(1 \leq i \leq q, 1 \leq j \leq k)$, $(q \geq p)$; and

(viii) $T = \mathrm{tr}\{[UFVGP_1H_1'R:0]E_1\}$.

Then for fixed U, V, P_1, and R

(a)
$$\sup_{F,G,H_1,E_1} T = \sum_{i=1}^{k} u_i v_i \rho_i r_i \ .$$

Furthermore, if

(ix) the u_i's, v_i's, ρ_i's, and r_i's are distinct and nonzero,

then

(b)
$$T = \sum_{i=1}^{k} u_i v_i \rho_i r_i$$

if and only if $G = \mathrm{diag}(\pm 1,\ldots,\pm 1)$, $F = \mathrm{diag}(\pm 1,\ldots,\pm 1)$,

$$H_1 = \begin{bmatrix} \pm 1 & . & 0 \\ 0 & \cdot & \pm 1 \\ \cdot & \cdot & \cdot \\ & 0 & \end{bmatrix} \quad \text{and} \quad E_1 = \begin{bmatrix} \pm 1 & . & 0 \\ 0 & \cdot & \pm 1 \\ \cdot & \cdot & \cdot \\ & 0 & \end{bmatrix} \quad \text{where}$$

G, F, H_1, and E_1 satisfy

$$g_{ii} f_{ii} h_{ii} e_{ii} = 1 \qquad (1 \leq i \leq k) \ .$$

Proof--

The theorem follows from a repeated application of Theorem 1. Using the notation of that theorem we have

$A = \{(G,F,H_1,E_1): G,F \in O(k),\ H_1 \in V(k,p),$ and $E_1 \in V(k,q)\}$,

$\alpha = (G,F,H_1,E_1)$, and

$B(\alpha) = [FVGP_1H_1'R:0]E_1$.

Let

$$\alpha_o = \left(I_k, I_k, \begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix}, \begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix} \right)$$

where I_k is the $k \times k$ identity matrix.

Then

$$\sup_{G,F,H_1,E_1} T \geq \operatorname{tr}\{UB(\alpha_o)\} = \sum_{i=1}^{k} u_i v_i \rho_i r_i \; .$$

Therefore to prove (a) it remains to prove that

$$(2) \qquad T \leq \sum_{i=1}^{k} u_i v_i \rho_i r_i \qquad \text{for all} \quad \alpha \in A \; .$$

If

$$(3) \qquad \sum_{i=1}^{\ell} b_{ii}(\alpha) \leq \sum_{i=1}^{\ell} v_i \rho_i r_i$$

for all $\alpha \in A$ and ℓ $(1 \leq \ell \leq k)$, then (2) follows from Theorem 1.

In the subsequent argument it will be necessary to consider various submatrices of G, F, H_1, and E_1. For any $r \times s$ matrix X let X^{mn} denote the $m \times n$ submatrix $(m \leq r, n \leq s)$ formed by the elements in the first m rows and n columns of X, i.e., if $X = (x_{ij})$ $(1 \leq i \leq r, 1 \leq j \leq s)$ then $X^{mn} = (x_{ij})$ $(1 \leq i \leq m, 1 \leq j \leq n)$.

Using this notation we have

$$\begin{aligned} b_{11}(\alpha) + \cdots + b_{\ell\ell}(\alpha) &= \operatorname{tr}\{F^{\ell k}[VGP_1H_1'R:0]E_1^{q\ell}\} \\ &= \operatorname{tr}\{[VGP_1H_1'R:0]E_1^{q\ell}F^{\ell k}\} \\ &= \operatorname{tr}\{VGP_1H_1'RE_1^{p\ell}F^{\ell k}\} \\ &= \operatorname{tr}\{VC(\ell)\} \\ &= \sum_{i=1}^{k} v_i c_{ii}(\ell) \end{aligned}$$

where $C(\ell) = (c_{ij}(\ell))$ is the $k \times k$ matrix

$$C(\ell) = GP_1 H_1' RE_1{}^{p\ell} F^{\ell k} .$$

In fact $C(\ell)$ also depends on α but we have suppressed the α to simplify the notation.

If

$$\sum_{i=1}^{m} c_{ii}(\ell) \leq \sum_{i=1}^{\min(m,\ell)} \rho_i r_i \tag{4}$$

for all $\alpha \in A$ and m $(1 \leq m \leq k)$, then (3) follows from (4) by Theorem 1. We have

$$\begin{aligned} c_{11}(\ell) + \cdots + c_{mm}(\ell) &= \mathrm{tr}\{G^{mk} P_1 H_1' RE_1{}^{p\ell} F^{\ell m}\} \\ &= \mathrm{tr}\{P_1 H_1' RE_1{}^{p\ell} F^{\ell m} G^{mk}\} \\ &= \mathrm{tr}\{P_1 D(\ell, m)\} \\ &= \sum_{i=1}^{k} \rho_i d_{ii}(\ell, m) \end{aligned}$$

where $D(\ell,m) = (d_{ij}(\ell,m))$ is the $k \times k$ matrix

$$D(\ell,m) = H_1' RE_1{}^{p\ell} F^{\ell m} G^{mk}$$

If

$$\sum_{i=1}^{n} d_{ii}(\ell,m) \leq \sum_{i=1}^{\min(n,m,\ell)} r_i \tag{5}$$

for all $\alpha \in A$ and n $(1 \leq n \leq k)$, then (4) follows from (5) by Theorem 1. We have

$$\begin{aligned} d_{11}(\ell,m) + \cdots + d_{nn}(\ell,m) &= \mathrm{tr}\{(H_1{}^{pn})' RE_1{}^{p\ell} F^{\ell m} G^{mn}\} \\ &= \mathrm{tr}\{RE_1{}^{p\ell} F^{\ell m} G^{mn} (H_1{}^{pn})'\} \end{aligned}$$

$$= \mathrm{tr}\{RS(\ell,m,n)\}$$

$$= \sum_{i=1}^{p} r_i s_{ii}(\ell,m,n)$$

where $S(\ell,m,n) = (s_{ij}(\ell,m,n))$ is the $p \times p$ matrix

$$S(\ell,m,n) = E_1{}^{p\ell} F^{\ell m} G^{mn} (H_1{}^{pn})' \ .$$

If

$$\sum_{i=1}^{q} s_{ii}(\ell,m,n) \leq \min(\ell,m,n,q) \tag{6}$$

for all $\alpha \in A$ and q $(1 \leq q \leq p)$, then (5) follows from (6) by Theorem 1.

Condition (6) is equivalent to the following two conditions. For all $\alpha \in A$

$$\mathrm{tr}\{S(\ell,m,n)\} \leq \min(\ell,m,n) \tag{7}$$

and

$$s_{ii}(\ell,m,n) \leq 1 \qquad (1 \leq i \leq p) \tag{8}$$

First consider (7). There are three possibilities. Namely, $\min(\ell,m,n) = \ell$, $\min(\ell,m,n) = m$, and $\min(\ell,m,n) = n$. Because the proof is essentially the same in all three cases, we assume that $\min(\ell,m,n) = \ell$. Then

$$\mathrm{tr}\{S(\ell,m,n)\} = \mathrm{tr}\{F^{\ell m} G^{mn} (H_1{}^{pn})' E_1{}^{p\ell}\}$$

$$= \sum_{i=1}^{\ell} f_i' G^{mn} (H_1{}^{pn})' e_i$$

where f_i' is the i-th row of $F^{\ell m}$ and e_i is the i-th column of $E_1{}^{p\ell}$. Since $F^{\ell m}$ and $E_1{}^{p\ell}$ are submatrices of orthogonal matrices, $f_i' f_i \leq 1$ and $e_i' e_i \leq 1$. Let $\alpha_i' = f_i' G^{mn}$ and $\beta_i = (H_1{}^{pn})' e_i$. By Lemma 1 $\alpha_i' \alpha_i \leq 1$ and $\beta_i' \beta_i \leq 1$. Hence, by the Cauchy-Schwarz inequality,

$$\mathrm{tr}\{S(\ell,m,n)\} = \sum_{i=1}^{\ell} \alpha_i'\beta_i$$

$$\leq \sum_{i=1}^{\ell} (\alpha_i'\alpha_i)^{\frac{1}{2}}(\beta_i'\beta_i)^{\frac{1}{2}}$$

$$\leq \ell = \min(\ell,m,n)$$

proving (7).

To prove (8) note that

$$s_{ii}(\ell,m,n) = e_i'F^{\ell m}G^{mn}h_i$$

where now e_i' is the i-th row of $E_1^{p\ell}$ and h_i is the i-th column of $(H_1^{pn})'$. Since $E_1^{p\ell}$ and H_1^{pn} are submatrices of orthogonal matrices, $e_i'e_i \leq 1$ and $h_i'h_i \leq 1$. Let $\alpha_i' = e_i'F^{\ell m}$ and $\beta_i = G^{mn}h_i$. By Lemma 1 $\alpha_i'\alpha_i \leq 1$ and $\beta_i'\beta_i \leq 1$. Hence, by the Cauchy-Schwarz inequality,

$$s_{ii}(\ell,m,n) = \alpha_i'\beta_i$$

$$\leq (\alpha_i'\alpha_i)^{\frac{1}{2}}(\beta_i\beta_i)^{\frac{1}{2}}$$

$$\leq 1$$

proving (8).

Let 'a → b' mean 'a implies b'. The preceding discussion can be summarized as follows. We have proved that (7) and (8) hold for all $\alpha \in A$ and that (7) and (8) → (6) for all q $(1 \leq q \leq p)$ → (5) for all n $(1 \leq n \leq k)$ → (4) for all m $(1 \leq m \leq k)$ → (3) for all ℓ $(1 \leq \ell \leq k)$ → (2). Therefore

$$T \leq \sum_{i=1}^{k} u_i v_i \rho_i r_i$$

for all $\alpha \in A$, proving (a).

Now assume that the u_i's, v_i's, ρ_i's, and r_i's are distinct and non-zero. We could prove (b) by a more careful analysis of the argument used to prove (a). In fact it is easier to prove (b) directly. Recall that

$$B(\alpha) = [FVGP_1H_1'R:0]E_1 .$$

If

(9) $$F = \mathrm{diag}(\pm 1,\ldots,\pm 1) , \quad G = \mathrm{diag}(\pm 1,\ldots,\pm 1) ,$$

$$H_1 = \begin{bmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \\ \cdot & \cdot & \cdot \\ & 0 & \end{bmatrix} \quad \text{and} \quad E_1 = \begin{bmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \\ \cdot & \cdot & \cdot \\ & 0 & \end{bmatrix}$$

where $f_{ii}g_{ii}h_{ii}e_{ii} = 1 \quad (1 \le i \le k)$, then

$$UB(\alpha) = \mathrm{diag}(u_1v_1r_1\rho_1,\ldots,u_kv_kr_k\rho_k)$$

and

$$T = \sum_{i=1}^{k} u_iv_ir_i\rho_i .$$

Conversely, suppose that

$$T = \sum_{i=1}^{k} u_iv_ir_i\rho_i .$$

It follows from Theorem 1 and (a) that

$$b_{11}(\alpha) = v_1r_1\rho_1$$

(10) $$b_{11}(\alpha) + b_{22}(\alpha) = v_1r_1\rho_1 + v_2r_2\rho_2$$

$$\vdots$$

$$b_{11}(\alpha) + \cdots + b_{kk}(\alpha) = v_1r_1\rho_1 + \cdots + v_kr_k\rho_k .$$

Let $\xi' = (f_{11}v_1,\ldots,f_{1k}v_k)$ and $\eta' = (e_{11}r_1,\ldots,e_{p1}r_p)$. Then from the Cauchy-Schwarz inequality

(11) $$b_{11}(\alpha) = \xi'GP_1H_1'\eta \le (\xi'\xi)^{\frac{1}{2}}(\eta'H_1P_1{}^2H_1'\eta)^{\frac{1}{2}}$$

with equality only if $\xi = cGP_1H_1'\eta$ for some constant c. Since $F \in O(k)$

(12) $$\xi'\xi = \sum_{i=1}^{k} v_i^2 f_{1i}^2 \leq v_1^2$$

with equality if and only if $f_{11} = \pm 1$. Similarly, since $E_1 \in V(k,q)$,

(13) $$\eta'\eta \leq r_1^2$$

with equality if and only if $e_{11} = \pm 1$. Let $\lambda = H_1'\eta$. Then from Lemma 2.3.1

(14) $$\lambda' P_1^2 \lambda \leq \rho_1^2 \lambda'\lambda$$

with equality if and only if $\lambda_i = 0$ $(2 \leq i \leq k)$. From Lemma 1

(15) $$\lambda'\lambda = \eta' H_1 H_1' \eta \leq \eta'\eta .$$

Combining (11)-(15) we have

$$b_{11}(\alpha) \leq v_1 r_1 \rho_1$$

with equality if and only if

$$G = \mathrm{diag}(\pm 1, G_2) , \quad F = \mathrm{diag}(\pm 1, F_2) ,$$

$$H_1 = \begin{bmatrix} \pm 1 & 0 \\ 0 & H_2 \end{bmatrix} , \quad E_1 = \begin{bmatrix} \pm 1 & 0 \\ 0 & E_2 \end{bmatrix}$$

where $f_{11}g_{11}h_{11}e_{11} = 1$, $G_2, F_2 \in O(k-1)$, $H_2 \in V(k-1,p-1)$, and $E_2 \in V(k-1,q-1)$. It follows from (10) that $b_{22}(\alpha) = v_2 r_2 \rho_2$. We can now use the preceding argument with G_2, F_2, H_2, and E_2 to show that $f_{22} = \pm 1$, $g_{22} = \pm 1$, $h_{22} = \pm 1$, and $e_{22} = \pm 1$ where $f_{22}g_{22}h_{22}e_{22} = 1$. Continuing in this manner we finally have

$$T = \sum_{i=1}^{k} u_i v_i r_i \rho_i$$

implies F, G, H_1, and E_1 satisfy (9), proving (b).

Corollary 1--

Suppose that

(i) $V = \mathrm{diag}(v_1,\ldots,v_k)$ with $v_1 \geq v_2 \geq \cdots \geq v_k \geq 0$;

(ii) $\Theta_1 = \mathrm{diag}(\theta_1,\ldots,\theta_k)$ with $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_k \geq 0$;

(iii) $L = \mathrm{diag}(\ell_1,\ldots,\ell_p)$ with $\ell_1 \geq \ell_2 \geq \cdots \geq \ell_p \geq 0$, $(p \geq k)$;

(iv) $G \in O(k)$;

(v) $H_1 \in V(k,p)$;

(vi) $E_1 \in V(k,n_1)$, $(n_1 \geq p)$; and

(vii) $T = \mathrm{tr}\{[VG\Theta_1^{\frac{1}{2}}H_1'L^{\frac{1}{2}}:0]E_1\}$.

Then for fixed V, Θ_1, and L

$$\text{(a)} \quad \sup_{G,H_1,E_1} T = \sum_{i=1}^{k} v_i\theta_i^{\frac{1}{2}}\ell_i^{\frac{1}{2}} .$$

Furthermore, if

(viii) the v_i's, θ_i's, and ℓ_i's are distinct and nonzero, then

$$\text{(b)} \quad T = \sum_{i=1}^{k} v_i\theta_i^{\frac{1}{2}}\ell_i^{\frac{1}{2}}$$

if and only if $G = \mathrm{diag}(\pm 1,\ldots,\pm 1)$,

$$H_1 = \begin{bmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \\ & 0 & \end{bmatrix} \quad \text{and} \quad E_1 = \begin{bmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \\ & 0 & \end{bmatrix}$$

where $g_{ii}h_{ii}e_{ii} = 1$ $(1 \leq i \leq k)$.

Proof--

T may be expressed as

$$T = \mathrm{tr}\{[V^{\frac{1}{2}}FV^{\frac{1}{2}}G\Theta_1^{\frac{1}{2}}H_1'L^{\frac{1}{2}}:0]E_1\}$$

where $F = I_k$ the $k \times k$ identity matrix. Corollary 1 now follows directly from Theorem 2 by adding the constraint $F = I_k$.

BIBLIOGRAPHY

Anderson, G. A. (1965). An asymptotic expansion for the distribution of the latent roots of the estimated covariance matrix. Ann. Math. Statist., 36, 1153-1173.

Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis, John Wiley and Sons, New York.

Bartlett, M. S. (1938). Further aspects of the theory of multiple regression. Proc. Camb. Phil. Soc., 34, 33-40.

Bartlett, M. S. (1939). A note on tests of significance in multivariate analysis. Proc. Camb. Phil. Soc., 35, 180-185.

Bartlett, M. S. (1941). The statistical significance of canonical correlations. Biometrika, 32, 29-37.

Bartlett, M. S. (1947a). Multivariate analysis. J. R. Statist. Soc. (Suppl.), 9, 176-190.

Bartlett, M. S. (1947b). The general canonical correlation distribution. Ann. Math. Statist., 18, 1-17.

Bellman, R. (1960). Introduction to Matrix Analysis. McGraw-Hill, New York.

Box, G. E. P. (1949). A general distribution theory for a class of likelihood criteria. Biometrika, 36, 317-346.

Burrill, C. W. (1972). Measure, Integration, and Probability. McGraw-Hill, New York.

Chattopadhyay, A. K. and Pillai, K. C. S. (1973). Asymptotic expansions for the distributions of characteristic roots when the parameter matrix has several multiple roots. Multivariate Analysis III (P. R. Krishnaiah, ed.), Academic Press, New York, 117-127.

Chattopadhyay, A. K., Pillai, K. C. S., and Li, H. C. (1976). Maximization of an integral of a matrix function and asymptotic expansions of distributions of latent roots of two matrices. Ann. Statist., 4, 796-806.

Chikuse, Y. (1974). Asymptotic expansions for the distributions of the latent roots of two matrices in multivariate analysis. Ph.D. Dissertation, Yale University.

Constantine, A. G. and James, A. T. (1958). On the general canonical correlation coefficient. Ann. Math. Statist., 29, 1146-1161.

Constantine, A. G. (1963). Some non-central distribution problems in multivariate analysis. Ann. Math. Statist., 34, 1270-1285.

Constantine, A. G. and Muirhead, R. J. (1972). Partial differential equations for hypergeometric functions of two argument matrices. J. Mult. Anal., 2, 332-338.

Constantine, A. G. and Muirhead, R. J. (1976). Asymptotic expansions for distributions of latent roots in multivariate analysis. J. Mult. Anal., 6, 369-391.

Deemer, W. L. and Olkin, I. (1951). The Jacobians of certain matrix transformations useful in multivariate analysis. Biometrika, 38, 345-367.

Erdélyi, A. (1956). Asymptotic Expansions. Dover, New York.

Fisher, R. A. (1939). The sampling distribution of some statistics obtained from non-linear equations. Ann. Eugenics, 9, 238-249.

Fujikoshi, Y. (1976). Asymptotic expansions for the distributions of the latent roots in MANOVA. (submitted for publication).

Goffman, C. (1965). Calculus of Several Variables. Harper and Row, New York.

Goursat, E. (1917). Differential Equations. Ginn and Company, New York.

Herz, C. S. (1955). Bessel functions of matrix argument. Ann. Math., 61, 474-523.

Hotelling, H. (1936). Relations between two sets of variates. Biometrika, 28, 321-377.

Hsu, L. C. (1948). A theorem on the asymptotic behavior of a multiple integral. Duke Math. J., 15, 623-632.

Hsu, P. L. (1939). On the distribution of the roots of certain determinantal equations. Ann. Eugenics, 9, 250-258.

Hsu, P. L. (1941a). On the limiting distribution of roots of a determinantal equation. J. Lond. Math. Soc., 16, 183-194.

Hsu, P. L. (1941b). On the limiting distribution of the canonical correlations. Biometrika, 32, 38-45.

James, A. T. (1954). Normal multivariate analysis and the orthogonal group. Ann. Math. Statist., 25, 40-75.

James, A. T. (1961). Zonal polynomials of the real positive definite symmetric matrices. Ann. Math., 74, 456-469.

James, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. Ann. Math. Statist., 35, 475-501.

James, A. T. (1966). Inference on latent roots by calculation of hypergeometric functions of matrix argument. Multivariate Analysis (P. R. Krishnaiah, ed.) Academic Press, New York, 209-235.

James, A. T. (1968). Calculation of Zonal Polynomial Coefficients by Use of the Laplace-Beltrami Operator. Ann. Math. Statist., 39, 1711-1718.

James, A. T. (1969). Tests of equality of latent roots of the covariance matrix. Multivariate Analysis II (P. R. Krishnaiah, ed.). Academic Press, New York, 205-218.

Kshirsagar, A. M. (1972). Multivariate Analysis. Marcel Dekker, New York.

Lawley, D. N. (1959). Tests of significance in canonical analysis. Biometrika, 46, 59-66.

Luke, Y. L. (1969). The Special Functions and Their Approximations. Volume I. Academic Press, New York.

Magnus, W., Oberhettinger, F. and Soni, R. P. (1966). Formulas and Theorems for the Special Functions of Mathematical Physics. Springer-Verlag New York, Inc.

Marriot, F. H. C. (1952). Tests of significance in canonical analysis. Biometrika, 39, 56-65.

Muirhead, R. J. (1970). Systems of partial differential equations for hypergeometric functions of matrix argument. Ann. Math. Statist., 41, 991-1001.

Murnaghan, F. D. (1938). The Theory of Group Representations. The Johns Hopkins Press, Baltimore.

Rao, C. R. (1973). Linear Statistical Inference and Its Applications. John Wiley and Sons, New York.

Roy, S. N. (1939). p-statistics, or some generalisations in analysis of variance appropriate to multivariate problems. Sankhya, 4, 381-396.

Sugiura, N. (1976). Asymptotic expansions of the distributions of the canonical correlations. Commemoration Volume of the Founding of the Faculty of Integrated Arts and Sciences, Hiroshima University.

Wilks, S. S. (1932). Certain generalizations in the analysis of variance. _Biometrika_, _24_, 471-494.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 43	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ASYMPTOTIC DISTRIBUTIONS OF LATENT ROOTS IN CANONICAL CORRELATION ANALYSIS AND IN DISCRIMINANT ANALYSIS WITH APPLICATIONS TO TESTING AND ESTIMATION.		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) William J. Glynn		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0563
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Yale University New Haven, CT 06520		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 042-242
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics and Probability Program (Code 436) Arlington, VA 22217		12. REPORT DATE June, 1977
		13. NUMBER OF PAGES 145
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Asymptotic expansion
Canonical correlations
Latent roots
Discriminant analysis

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The principal aim of this research is to derive asymptotic expansions for the distributions of latent roots which occur in canonical correlation analysis and in discriminant analysis. The results are used to obtain estimates of the population roots and to study tests that some of the population roots are zero.

(Over)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Suppose that $\rho_1^2 > \rho_2^2 > \cdots > \rho_k^2 > \rho_{k+1}^2 = \cdots = \rho_p^2 = 0$ are the squared canonical correlation coefficients between a p-variate random vector x and a q-variate $(q \geq p)$ random vector y where x and y have a joint normal distribution. Let $r_1^2,\ldots,r_p^2$ be the corresponding maximum likelihood estimates of the ρ_i^2's based on a random sample of size $N = n + 1$. An asymptotic expansion for the joint density function of the r_i^2's for large n is obtained. The expansion is based on an asymptotic expansion of the ${}_2F_1$ hypergeometric function of two matrix arguments which occurs in the exact joint density of the r_i^2's. The expansion is used to obtain marginal maximum likelihood estimates of the ρ_i's and to study the Bartlett-Lawley tests that the residual population coefficients are zero.

Let $\ell_1 > \ell_2 > \cdots > \ell_k > \ell_{k+1} = \cdots = \ell_p = 0$ be the latent roots of the matrix $B(B+W)^{-1}$ where W is a $p \times p$ random matrix having a Wishart distribution $W_p(n_2,\Sigma)$ and B is a $p \times p$ random matrix having a noncentral Wishart distribution $W_p(n_1,\Sigma,\Omega)$. In discriminant analysis W and B are respectively the "within groups" and "between groups" matrices of sums of squares and sums of products. An asymptotic expansion for the joint density function of the ℓ_i's for large n_2 is obtained under the assumption that Ω is of the form $\Omega = n_2\Theta$. The expansion is based on an asymptotic expansion of the ${}_1F_1$ hypergeometric function which occurs in the exact joint density of the ℓ_i's. The expansion is used to study Bartlett's test that the residual θ_i's are zero.